Application of molecular imaging to address current clinical challenges in breast cancer

Glendenning, Jennifer Louise

Awarding institution:
King's College London

Application of molecular imaging to address current clinical challenges in breast cancer

Dr Jennifer Glendenning

MD(Res) Thesis

July 2016

1 Table of Contents

1.1 List of tables

1.2 List of figures

1.3 List of abbreviations

A

ANC	Axillary Nodal Clearance
ATTC	American Type Tissue Culture Collection

B

BLBC	Basal-like breast cancer
BLI	Bioluminescent imaging
BSA	Bovine serum albumin
BSC	Breast Conserving Surgery

C

cTIMP	Clinical trial of investigational medicinal product
C.I.	Confidence Interval
^{64}Cu	Copper 64 isotope

D

DARPin	Designed Ankyrin Repeat Protein
DCE-MRI	Dynamic contrast enhanced MRI
DFS	Disease Free Survival
DMEM	Dulbecco's modified Eagles medium
DMSO	Dimethyl sulphoxide

E

EANM	European Association of Nuclear Medicine
eCRF	Electronic case report format
EFS	Event free survival
ER	Oestrogen receptor
EORTC	European Organization for Research and Treatment of Cancer

F

FBS	Fetal Bovine Serum
FDG	2-[fluorine18]fluoro-2-deoxy-D-glucose, [^{18}F]FDG
FLT	3'deoxy-3'-[fluorine18]fluorothymidine, [^{18}F]FLT
FTV	Functional Tumour Volume

G

^{68}Ga	Gallium 68 isotope
GLUT-1	Glucose Transporter 1
GTV	Gross Tumour Volume

H

H&E	Haematoxylin and eosin

HER2	Human epidermal growth factor receptor 2
HER2+	Human epidermal growth factor receptor 2 positive
HTA	Human Tissue Authority
I	
IB	Investigator Brochure
ICR-CTSU	Institute of cancer research clinical trials and statistics unit
IHC	Immunohistochemistry
^{111}In	Indium 111 isotope
IMP	Investigational Medicinal Product
IMPD	Investigational Medicinal Product Dossier
IRAS	Integrated research application service
IV	Intra venous
K	
k1	PET tracer transport rate from arterial plasma into displaceable tissue compartment, ml/min/g
Ki	PET tracer metabolism influx, ml/min/g
KHP-CTO	Kings Health Partners Clinical Trials Office
L	
LBM	Lean Body Mass
M	
M	Mastectomy
μm	Micron
MRI	Magnetic Resonance Imaging
N	
nM	nanomolar
NST	Invasive breast cancer no special type
O	
OS	Overall survival
P	
PBS	Phosphate saline buffer
pCR	Pathological Complete Response
PET	Positron Emission Tomography
PERCIST	PET Response Criteria in Solid Tumors
PIS	Patient information sheet
PR	Progesterone receptor
R	

RCB	Residual Cancer Burden
REC	Research Ethics Committee
RECIST	Response Evaluation Criteria in Solid Tumours
RFB	Red fluorescent protein
RFS	Relapse Free Survival
ROI	Region of Interest
RPM	Revolutions per minute
S	
SC	Subcutaneous
s.d.	Standard deviation
SEM	Standard error of the mean
SNB	Sentinel lymph node biopsy
SPECT	Single-photon emission computed tomography
SUV	Standardised Uptake Variable
SUL	Lean body mass corrected Standardised Uptake Variable
T	
TBST	Tris-Buffered Saline and Tween 20
TNBC	Triple negative breast cancer
t½	radio-isotope half life
Z	
^{89}Zr	Zirconium 89 isotope

1.4 Acknowledgements

I am indebted to the many clinical colleagues working within the Breast Unit at Guys and St Thomas's NHS Foundation Trust and Kings College Hospital NHS Foundation Trust who kindly agreed to allow me to recruit patients to the TNPET01 study, in particular oncology colleagues Dr Anna Rigg, Dr Mark Harries and Dr Hartmut Kristeleit; Dr Jyoti Parikh, Mrs Julie Scudder and Professor Vicky Goh from the radiology team; and to Dr Julie Owen and Dr Natalie Woodman Dr Patrycja Gazinska (PG) and Professor Sarah Pinder (SP) in research pathology. The research PET imaging could not have happened without fantastic support from the team at the PET Centre, especially the clinical scientists Dr Joel Dunn, and Dr Jim O'Doherty and Dr Lucy Pike; Dr Kam Kahlon in radiochemistry for persisting with an unwell cyclotron and the imaging technologists Jon Joemon and Armidita Jacob who somehow managed to find research imaging slots despite the pressures of clinical work. I am grateful to Professor Judith Bliss, Dr Holly Tovey and Dr Lucy Kilburn from the ICR-CTSU for the statistical support provided freely for the TNPET01 study.

I would like to thank the Breakthrough Breast Cancer Unit, KCL, in particular Dr Rebecca Marlow, Dr Daniel Weekes, fellow research student Nirmesh Patel and Breakthrough laboratory technician, Erika Francesch Domenech. Without their patient training, encouragement and delicious cakes the preclinical work could not have been undertaken. I am grateful to Dr Robert Goldstein (UCL) who provided the G3 DARPin, and to the Imaging Sciences team especially Dr Margaret Cooper (KCL), Dr Gilbert Fruhwirth and Dr Florian Kampermeir who undertook the radiolabelling and freely shared their technical expertise and advice in the face of adversity.

I would like to thank the Comprehensive Cancer Imaging Centre, Breakthrough Breast Cancer and the Guys and St Thomas' Charity who provided me with financial support to undertake this work; and my supervisors Professor Andrew Tutt, Professor Michael O'Doherty and Dr Sally Barrington; and my family, Tim, Ellie and Tom who supported me more than they realise.

Finally, but most importantly, to the patients who freely gave their time and energy to participate in the TNPET01 study at a very difficult time in their lives.

1.5 Publications and Presentations arising from this work

Glendenning J, Barrington S, Tovey H, Dunn J, Tutt A. Repeatability evaluation of PET/CT imaging using [18F]fluorothymidine (FLT) and [18F]fluorodeoxyglucose (FDG) in primary triple negative breast cancer (TNBC). IMPAKT Brussels May 2015 abstract #210

Glendenning J, TNPE01-phase II trial of FLT and FDG PET vs. MRI to assess therapy response of breast cancer. Imaging in Stratified Cancer Treatment-Methodology, Preclinical Discovery and Clinical Trials. Newcastle, Nov 2013 (winning best oral presentation)

Glendenning J, Cook G. Imaging Breast Cancer Bone Metastases: Current Status and Future Directions. Seminars Nucl Med 2013; 43(4):317-323

Glendenning J, Irshad S, Tutt A. Treatment of Metastatic Triple-Negative Breast Cancer. Current Breast Cancer Reports. 2012; 4(1):10-21

Glendenning J, Tutt A. N. PARP inhibitors – current status and the walk towards early breast cancer. 2011. The Breast; 20(3):S12-S19

2 Abstract

Breast cancer has multiple sub-types, designated in clinical practice by the presence or absence of prognostic and predictive biomarkers: oestrogen receptor, progesterone receptor and the human epidermal growth receptor 2 (HER2). Molecular imaging offers the opportunity firstly to non-invasively and dynamically interrogate *in vivo* tumour sites to provide prognostic and predictive evaluation at clinically meaningful on-treatment time points. Secondly, it may address diagnostic challenges in the metastatic setting by enabling assessment of inter- and intra-lesion heterogeneity without requirement for tissue acquisition from multiple sites.

This thesis describes pre-clinical evaluation of a novel HER2 targeted DARPin radiotracer as an *in vivo* diagnostic across multiple tumour sites in murine models. Metastatic breast xenograft models were developed and validated using bioluminescence imaging and definitive histology for *in vivo* evaluation of a novel HER2 targeted DARPin radiotracer. In subsequent preclinical testing the DARPin radio-tracer failed to differentiate HER2 status of pre-clinical tumour xenografts models and this data raises significant questions regarding suitability of the DARPin radiotracer for clinical evaluation as a HER2 diagnostic.

Additionally this work reports the set-up of a Phase 2 imaging feasibility study designed in two parts to evaluate post-cycle 1 PET response using the FLT- and FDG-PET imaging tracers to address clinical questions concerning tracer selection, scan acquisition and interpretation for validation of Positron Emission Tomography (PET) response as a predictive biomarker of neoadjuvant response in the triple negative breast cancer (TNBC) phenotype. Part A (participant recruitment completed) delivers the first phenotype specific repeatability constraints for the most commonly reported standardised uptake parameters (SUV); maximum (SUVmax), mean (SUVmean), peak (SUVpeak) and lean body mass corrected peak (SULpeak), assessed at conventional (90 minutes) and exploratory (120 and 180 minute) acquisition time points. The TNBC SUV intrinsic variability was 12-24% in both tracers and is dependent on scan acquisition time and SUV parameter. The FDG tracer has progressed to the second phase, Part B, to provide the first TNBC phenotype specific response data at a post-cycle 1 time point. The data suggests SUV change can predict later residual cancer burden and that >40% threshold change will be required to differentiate RCB 0-1 vs 2-3 response, a change that exceeds current EORTC/PERCIST recommendations for solid tumour chemotherapy response prediction. The study will inform future use of early FDG-PET as an exploratory biomarker in window of opportunity and novel therapy neo-adjuvant trials in TNBC

3 Background literature review

3.1 Introduction

Breast cancer is the commonest female malignancy accounting for almost one third of cancers (1). About 30% will relapse and it is second only to lung cancer as the leading cause of female cancer death. Breast cancer has multiple sub-types, designated in clinical practice by presence or absence of oestrogen receptor (ER), progesterone receptor (PR) and the human epidermal growth receptor 2 receptor (HER2). Better understanding of breast cancer biology and development of associated targeted therapies has driven improvement in outcomes particularly in hormone receptor positive and HER2 positive disease and guides therapy in these phenotypes (2, 3).

Triple negative breast cancer (TNBC) describes the 15-20% of breast cancers that are hormone and HER2 receptor negative. Clinical challenges in TNBC arise from the lack of targeted therapy and heterogeneity of response to standard cytotoxic therapy. Subgroups are highly sensitive to neoadjuvant chemotherapy achieving pathological response (pCR) and excellent long-term outcomes following sequential anthracycline-taxane regimens. However response is heterogeneous and overall TNBC is characterised by higher rates of distant recurrence than other phenotypes (4, 5) and little durable benefit from therapy at relapse (6). Molecular imaging shows promise as a predictive factor for chemosensitivity but there remains a need for validated surrogates of later pathological response and survival outcomes to clearly differentiate populations unlikely to derive benefit from standard neoadjuvant chemotherapy early in the course of treatment.

HER2 receptor amplification, present in 20-30% of breast cancers, has prognostic significance identifying a population of women at relatively increased risk of metastatic visceral and cerebral relapse (7, 8) which can be beneficially impacted by HER2 targeted therapy. HER2 evaluation is a mandatory part of the pathological assessment of new primary breast cancer, informing neoadjuvant and adjuvant use of HER2 targeted therapeutics such as Trastuzumab. In metastatic disease, evolution of HER2 molecular targeted therapies now offers clinicians a range of HER2 targeted options and coordinated transition though approved agents at radiological progression has improved survival for those with advanced disease over the last decade (9). Whilst pathological HER2 assessment is straightforward in primary disease, practical issues associated with tissue acquisition from metastatic sites frequently preclude re-evaluation of HER2 status. Consequent the clinical assumption of constancy in gene expression from primary disease through to metastasis diagnosis has significant therapeutic implications

and risks futile use or inappropriate omission of HER2 targeted agents for some individuals (10, 11).

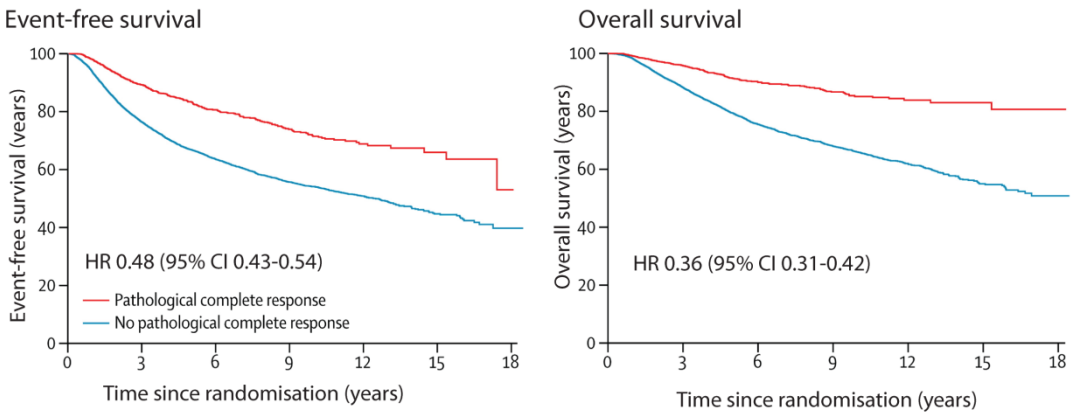
Molecular imaging offers the opportunity to non-invasively interrogate tumour sites at baseline and during the course of therapy, potentially providing prognostic and predictive evaluation at clinically meaningful diagnostic and on treatment time points. Developing validated imaging biomarker strategies offer opportunities to address specific clinical challenges posed by the different breast cancer subgroups, specifically the ability for timely on treatment response assessment and phenotype evaluation across a totality of metastatic disease sites. This would then facilitate individualised therapeutic selection unfettered by the practical or time point constraints inherent to current tissue diagnostic and cross-sectional imaging response evaluations relied upon in current clinical practice.

3.2 Molecular imaging as a predictive response biomarker in TNBC

3.2.1 Rationale for monitoring neoadjuvant systemic therapy response in TNBC

Curative treatment for locally advanced or inflammatory breast cancer increasingly includes systemic therapy given in the neoadjuvant setting to facilitate breast conservation whist delivering at least the same magnitude of survival advantage as the same adjuvant therapy (12-14). Pre-surgical sequencing also uniquely enables assessment of *in vivo* sensitivity to therapy. The absence of residual invasive tumour within breast or axilla at definitive surgery (pathological complete response, pCR) independently predicts better disease free and overall survival regardless of breast cancer subtype (Figure 3.1 A). Pooled data from 12 international neoadjuvant trials confirms this predictive impact of pCR is driven by the TNBC (EFS: HR 0·24, 95% CI 0·18–0·33; OS: 0·16, 0·11–0·25) and the trastuzumab treated HER2-positive (HER+), hormone-receptor-negative tumour (EFS: 0·15, 0·09–0·27; OS: 0·08, 0·03, *0*·22) subsets compared to hormone receptor positive disease (15).

A Associations between pCR and event-free and overall survival



B Frequency and estimate 10year RFS of RCB classes in breast cancer subsets

	TNBC		HER2+		ER+/HER2-	
RCB class	**Frequency**	**RFS (95% CI)**	**Frequency**	**RFS (95% CI)**	**Frequency**	**RFS (95% CI)**
pCR	35%	82% (63-92)	36%	83% (66-92)	10%	83% (59-93)
RCB-1	15%	82% (53-94)	16%	69% (41-86)	13%	96% (74-99)
RCB-2	33%	54% (38-68)	30%	50% (31-66)	60%	76% (68-83)
RCB-3	17%	18% (6-36)	18%	32% (13-52)	17%	45% (27-61)

Figure 3-1 Prognostic importance of pathological response on survival outcomes

A) Association between pathological complete response (pCR) and event-free survival (EFS) and overall survival. Responder analysis from 12 international trials including 11 955 patients. pCR defined as ypT0/is ypN0. HR=hazard ratio. Figure reproduced from Cortazar et al 2014 (15)

B) Frequency and estimated 10-year relapse free survival (RFS) of residual cancer burden (RCB) classes in breast cancer subsets following taxane-anthracycline neoadjuvant chemotherapy(16). In TNBC, RCB-II/III defines high-risk patient post chemotherapy subpopulations. Figure adapted from Symmans et al 2013 (16)

The prognostic insight provided by evaluation of pCR which predicts disease free survival (DFS) has the potential to more efficiently evaluate promising new agents in the early breast cancer context particularly in luminal B/HER2-negative, HER2-positive (non-luminal), and triple-negative disease (17)(14, 15). Recognising this, pCR has been the primary endpoint in recent neoadjuvant trials (18, 19) and accepted by the U.S. Food and Drug Administration (FDA) as a suitable surrogate efficacy endpoint leading to accelerated drug approval. However, the magnitude of increase in pCR rate required for a novel therapeutic to deliver meaningful rates of improved survival remains uncertain. Furthermore some categories of patients with incomplete response derive good long term outcomes from standard treatment implying that

dichotomous separation into pCR or not oversimplifies the prognosis for those with residual disease, ignoring primary tumour characteristics and magnitude of chemotherapy response (20, 21). With the aim of relating degree of partial response to subsequent survival outcome a spectrum of histopathological scoring systems have been described. These include binary pCR or not (NSABP-B18); linear histologic response in breast only (Miller-Payne(20)) or breast and lymph nodes (Sataloff (22)); algorithms integrating breast and nodal response including formula (Nottingham Prognostic Index (23)) or Web calculator (Residual Cancer Burden, RCB (21)); and, finally, those which integrate pretreatment clinical stage (C), ER status (E), and grade (G) with post-treatment pathological stage (PS) such as the CPS EG) (24) have been described. Overall, scoring systems which include amalgamated breast and nodal response better correlate with survival (25) and RCB is a recommended standard for use in neoadjuvant trials (26). RCB considers residual primary tumour dimension, cellularity and axillary nodal burden to calculate a score which can be categorised into four independently prognostic classes across all breast cancer phenotypes (21). Achieving RCB 0 or 1 (pCR or minimal residual tumour burden) predicts an excellent long term prognosis whereas RCB 2 and 3 (moderate and extensive residual tumour burden) identifies individuals at increasing risk of relapse, the magnitude of which is driven by the underlying breast cancer phenotype but is worst in the context of triple negative disease compared with other breast cancer subtypes (Figure 3.1 B) (16, 27).

Improving TNBC survival outcomes requires better understanding of heterogeneity in disease and therapy response. Tissue evaluations performed on the diagnostic biopsy specimen may indicate potential chemo-sensitivity but despite consistently greater pCR rates reported in triple negative, HER2+ breast cancer or in the molecularly defined basal-like and HER2+ subgroups compared to ER+ subgroups (5, 28) (29) the baseline features predictive of good response overlap with those predicting progression (30). At present prospective differentiation between later responding and RCB non-responding individuals remains an aspiration.

Neoadjuvant on treatment monitoring relies on clinical examination prior to each cycle and cross-sectional imaging performed at baseline, the mid-point of sequential therapy and prior to definitive surgery after 6-8 cycles of sequential taxane-anthracycline treatment. High spatial resolution contrast enhanced magnetic resonance imaging (CE-MRI) is the gold-standard imaging modality. The Response Evaluation Criteria in Solid Tumour (RECIST version 1.1) formally quantifies response by defining and monitoring change in evaluable disease using minimum unidimensional size criteria (31). Serial imaging classifies into four categories of response; complete (disappearance of all tumour foci sustained for at least 4 weeks); partial (decline of ≥30% in tumour diameter), stable (response that does not meet criteria for

progression nor response) and progressive (≥20% increase in size or new tumour lesions). Inevitably there is a time lag between therapy initiation and measurable size change and available data does not support MRI RECIST assessment earlier than the suggested interval of 6-8 weeks (32). Hence the first cross sectional imaging opportunity follows a minimum of 3 cycles of therapy. As this approaches scheduled transition to the second component of sequential taxane-anthracycline chemotherapy, RECIST provides little opportunity for therapy modification.

With the aim of facilitating earlier neoadjuvant response MRI assessment the multicenter Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging And moLecular Analysis (I-SPY TRIAL) investigated MRI- and tissue-based biomarkers for predicting response and survival (33, 34). For RCB prediction, the MRI derived functional tumour volume (FTV) was superior to clinical examination at all timepoints with the greatest relative benefit at following cycle 1 chemotherapy. However end of treatment MRI FTV best predicted RCB class, and equal contribution of MRI FTV following 2 cycles and the histopathological variables (RCB class and tumor subtype defined by hormone and HER2 receptor status) provided the strongest predictor of RFS (33). Further to this data, the currently recruiting I-SPY 2 trial (NCT NCT01042379) utilises an adaptive trial design which aims to integrate clinical, imaging including MRI function tumour volume response, and genomic data to identify agents for different molecular phenotypes early in the drug development cycle that improve pathologic response. Importantly early FTV does not independently predict later RCB or survival outcomes and is not suitable for adoption into routine clinical practice.

GeparTrio is the largest reported neoadjuvant breast cancer study to prospectively evaluate the concept of response adaptive chemotherapy (35). In this study the magnitude of USS response in the primary breast tumour following 2 cycles of combined anthracycline-taxane regimen defined randomisation to prolongation of the same regimen in the responding subset (further 6 vs. 4 cycles of the same regimen) or continuation vs. early chemotherapy transition to capecitibine/vinoerlbine in non-responders. Overall, disease-free survival was significantly longer in the USS response-guided groups than in the conventional groups combined (HR = 0.71, *P* = .0003) and exploratory subset analysis suggested greatest efficacy for response-guided therapy selection in hormone receptor–positive tumors. However the mix of phenotypes and unknown efficacy impact of transition to vinorelbine/capecitabine combination in poorly responding TNBC mean questions remain about the broader applicability of the strategy. Alternative approaches relying on significant residual disease to define high risk post neoadjuvant populations eligible for recruitment to adjuvant trials are possible but impractical for all but the least toxic adjuvant interventions given the impact of

the proceeding cytotoxic on fitness. Biopsy derived biomarkers such as ki-67 may have prognostic and predictive potential (36) and therapy induced changes in the tumour potentially correlate with subsequent survival outcome (37) and may add to RCB (38). However repeated invasive biopsy for *in vivo* response monitoring and potential confounding effect of intra-tumour heterogeneity (39) make serial tissue acquisition unattractive for routine clinical application. In contrast a validated imaging biomarker selected to reflect the biology of the disease in question and predictive of the pathological response surrogate may better realise opportunities afforded by the pre-surgical window for efficient evaluation of therapeutics. However without a validated means of early on-treatment monitoring in TNBC, pathological efficacy outcomes are only reached at therapy completion currently meaning that opportunities to tailor therapeutic strategy to *in vivo* tumour response are missed.

3.2.2 PET Functional Imaging for neoadjuvant response prediction in TNBC

Single Photon Emission Computerised Tomography (SPECT) and (Positron Emission Tomography (PET) molecular imaging are real time imaging techniques which map expression of molecular markers within humans or animals following the administration of radiolabelled tracers (40). Radioisotopes, which emit a single gamma photon, are detected following collimation to exclude those with a non-parallel trajectory and computer reconstruction of the spatial distribution of events generates 3D SPECT images. In contrast radioisotopes used in PET imaging emit positrons that annihilate after combination with an electron to generate two 180◦ opposing gamma photons. These are detected by paired coincidence coupled detectors, which locate the source of the annihilation event generating the spatial information required for computer reconstruction. For both modalities high contrast image acquisition requires a match between the physical radionuclide half-life and circulation half-life of the conjugated targeting ligand. Specific imaging requirements of targeting ligand are rapid delivery to the tumour site (efficient tissue delivery, cell penetration, high tumour specificity and high target affinity) and rapid excretion of unbound ligand to minimise the interval from tracer infusion to optimal image acquisition. Both modalities have an established role in clinical practice but PET has advantages of greater sensitivity, higher spatial resolution, and more accurate quantification.

The ability of PET to detect drug induced changes in tumour metabolism and proliferation that may anticipate size change has led to an emerging role informing escalation/de-escalation strategies within the research setting in highly proliferative malignancies such as lymphoma (e.g.NCT01304849, NCT00392314, NCT01361191). Given the importance of pathological response in TNBC, similar prospective trials testing early imaging response in adaptive

neoadjuvant chemotherapy approaches would be of clinical interest. However in a phenotype characterised by absence of molecular targets for imaging or therapy the most appropriate imaging biomarker remains unclear.

3.2.3 2-[fluorine18]fluoro-2-deoxy-D-glucose (FDG) as a candidate imaging biomarker

2-[fluorine18]fluoro-2-deoxy-D-glucose (FDG) is the most widely used PET tracer for staging and response assessment in cancer. Uptake reflects an increase in general cell metabolism requiring glucose as an energy source. The tracer is not tumour specific however increased expression of glucose transporters and hexokinase (glucose phosphorylating enzyme) as a result of oncogenic transformation leads to increased rates of glycolysis and glucose transport in most malignant tissues (41). Following intracellular phosphorylation, FDG becomes trapped and concentration within a tumour represents the glycolytic activity of viable malignant cells. The magnitude of FDG uptake could be impacted by effective therapy. Although potential false positive FDG uptake in areas of inflammation risks confound specificity for therapy (42), changes in biopsy derived measures of apoptosis (43) and proliferation (34) following cytotoxic therapy have both been shown to correlate with change in FDG uptake. Consequently, serial FDG PET imaging may indirectly provide important prognostic information about the viable tumour burden and anti-proliferative effect of cancer therapy.

Established biological factors predicting enhanced baseline FDG breast cancer uptake include high Ki-67 proliferative index, negative oestrogen receptor status and non-lobular histology (32, 44-49) (50). Differences in tumour blood flow, glycolysis and intracellular phosphorylation may account for heterogeneity of uptake (51-53). Expression of the facilitative glucose transporter 1 (GLUT-1) is particularly associated with enhanced FDG uptake (51) and overlap between genes associated with increased glucose metabolism and ER- molecular phenotypes has been reported (49). GLUT-1 expression characterised in 286 breast tumour samples classified as basal like breast cancer (BLBC) or otherwise using immunohistochemical (IHC) criteria reported significant association between GLUT-1 expression and BLBC (76.4% BLBCs vs. 23.8% non-BLBCs showing immunostaining for GLUT-1 , $p < .001$) (54).

3.2.3.1 FDG PET for TNBC response assessment

A literature search (Appendix 11.1) identified 32 published neoadjuvant FDG response evaluation studies. The majority (23 out of 33 studies) have considered FDG performance in mixed breast cancer phenotypes. Data interpretation is further complicated by heterogeneity of response definition, cytotoxic regimen, and timing of response evaluation, which may

account for the marked variability in published thresholds, sensitivities and specificities for pCR prediction. A 2012 meta-analysis (19 studies) supported the use of FDG-PET imaging early in the course of therapy (55) but relationships between phenotype, magnitude of SUV change and later response remained unclear. Subsequent to the meta-analysis a single study in a pre-specified TNBC population suggested <42% threshold change in SUVmax after 2 cycles defines a subpopulation at high risk of residual disease and early relapse (100% vs. 45% according to response; p= 0.014 and 44% vs. 0%; P = 0.024 respectively) (56). In a 2014 update of this study (n=50), the HR for relapse was 3.68 (95% CI 1.13-11.99) for patients with <42% decrease in SUVmax after 2 cycles anthracycline based chemotherapy (57). At the post cycle 1 time point a trend for the most marked SUV reductions has been observed in TNBC and HER2- populations but no data prospectively addresses the TNBC population specifically (58-61).

3.2.3.2 Areas of uncertainty in FDG breast cancer response assessment

3.2.3.2.1 Which reporting parameter should be used for serial FDG scan assessment?

The majority of breast cancer response evaluation studies (Appendix 11.1) used a single static scan to measure and report changes in the standardised uptake value (SUV). This is a semi-quantitative measure, which describes the ratio of tissue concentration and injected activity without considering tracer kinetics. Normalisation is usually to body mass but may also be to lean body mass (SUL) or body surface area. A variety of SUV parameters are reported within the published literature (Figure 3-3A) and comparative data is required to establish the most clinically meaningful predictive measure for differentiating FDG response .

SUVmax, the hottest pixel within the tumour volume (figure 3.3B), is the most widely reported parameter in oncology publications. SUVmax is easily quantified on commercial workstations and more resistant to partial volume effects in small tumours than other SUV parameters which may be advantageous in breast cancer. However wider concerns about adverse impact of imaging noise on quantification of cancer treatment response led to development of PERCIST guidance advocating the use of the peak parameter defined as the average SUV within a small, fixed-size region of interest (ROI_{peak}) centred on the highest uptake part of the tumour but not necessarily including the SUVmax pixel (62) (Figure 3-4A). Superiority of peak over max parameter is uncertain and in a predominantly HER2+ve and ER+ve breast cancer cohort two cycles of chemotherapy induced parallel therapy changes for the two parameters (32). The peak parameter has not been widely adopted and most (21 of 25 breast cancer FDG response studies) publications during or subsequent to 2009 continue to report therapy induced change in SUVmax (Appendix 11.1).

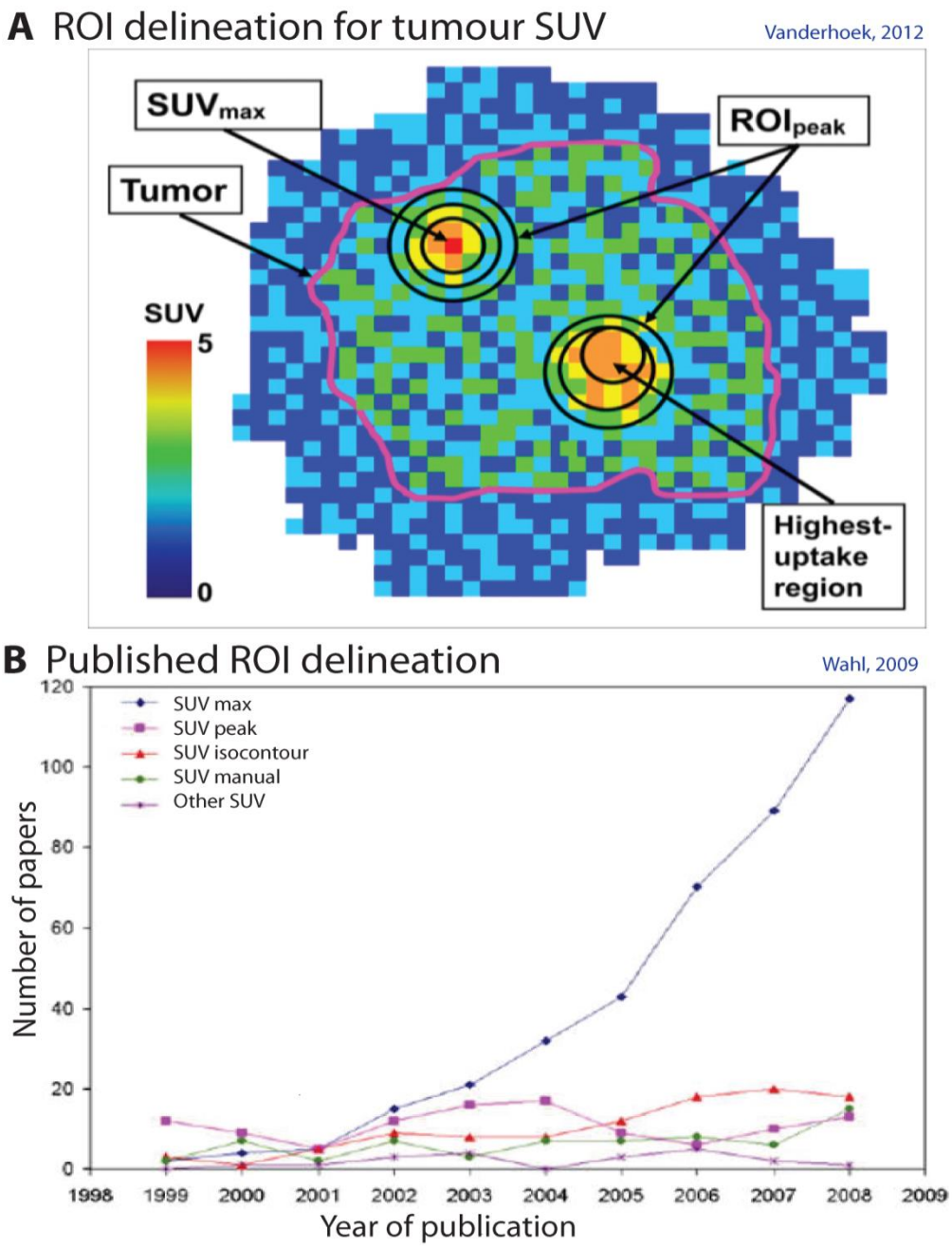


Figure 3-2 Schematic PET image of radiotracer uptake in tumour

A. Schematic PET image of radiotracer uptake in tumour (purple line) demonstrating impact of ROI definition on SUV parameters within a tumour volume (delineated in pink) (63). SUVmax refers to maximal pixel in tumour. SUVpeak is the average SUV within a small fixed-sized region of interest whose value will vary according to placement with the tumour. Image taken from Vanderhoek et al 2012(63).

B Number of human FDG oncology papers that included use of tumour SUV, as function of year of publication (62). SUVpeak and SUVmax as above. SUV isocontour refers to the mean SUV value across an irregular ROI defined by isocontour set at a percentage of the maximal pixel. SUV manual refers to the mean SUV across a manually drawn ROI. Only a subset of these papers describes response assessment studies. Image taken from Wahl et al 2009 (62).

The peak and max parameters provide an SUV estimate for the most metabolically active region of the tumour volume (figure 3.3A) but a mean SUV reporting glycolysis across the total tumour may potentially be advantageous for tumour response evaluation. In squamous cell carcinoma of the head or neck metabolically active tumour volumes (MTV) delineated using FDG PET are superior to those defined on CT or MRI when compared to reference volume assessed from surgical specimens, (64) and the isocontour at 40% of maximum SUV (SUV 40 isocontour) delivers the best compromise between accuracy and risk of underestimating tumour extent (65). Similar, tissue comparative data supports superiority of SUV40-42 isocontours in cervical and lung cancer (66, 67) but relationships with pathological size in breast cancer are unknown. PERCIST suggests SUVmean as an exploratory response parameter but acknowledges limited data and this parameter is reported in few breast cancer response publications (Appendix 11.1). In ER+ HER2- disease, Total Lesion Glycolysis (TLG, the product of metabolically active tumour volume and SUVmean) was reported to have better predictive value than SUVmax (68). However in a heterogeneous breast cancer population SUVmean (42% isocontour) conferred no advantage over SUVmax response at post cycle 1, midpoint or end of treatment timepoints. In 16 patients (19 lesions) with mixed metastatic solid primary tumours undergoing mid and end of treatment PET response assessment conflicting response categorisation was present in 80% of participants due to significant differences in magnitude of SUV change between the different SUV measures (max, mean, peak and total)(63). There is a lack of comparative data in TNBC but it is likely that classification of PET response will similarly be dependent on SUV parameter, supporting the need for clinical trials to select the most predictive SUV measure for assessment of therapy response.

3.2.3.2.2 Defining FDG PET response in TNBC?

Level of SUV threshold change to define responders will likely be influenced by the objective of PET imaging. The clinical aspiration is for early identification of high risk subpopulations in whom escalation or novel therapeutic approaches would be of interest. In this context the post cycle 1 time point has maximal potential to prevent ineffective therapy in non-responders. In breast cancer it is known that the greatest magnitude of change occurs after the cycle 1 chemotherapy (Figure 3-3) (69-73). The 2012 meta-analysis supported early FDG PET response assessment after the 1st or 2nd cycle of chemotherapy and suggested a discriminatory threshold of 55-65% but cautions that these data are from mixed breast cancer phenotypes with an unknown or minority TNBC proportion (55). Very early assessment (day 8 following first chemotherapy) risks confounding evaluation of SUV increase through detection of therapy instigated inflammatory response rather than active tumour (72). Therefore an

interval of at least 2 weeks between completion of the chemotherapy cycle and the [18F]-FDG PET scan is likely required to avoid transient increases.

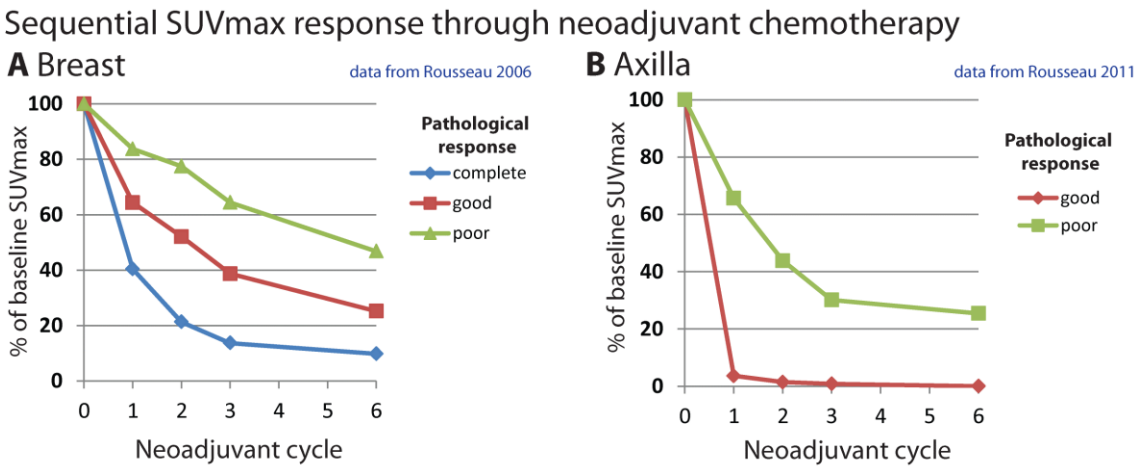


Figure 3-3 Sequential SUV change in breast and axillary lesions through neoadjuvant chemotherapy

A Figure adapted from adapted from Rousseau et al 2006 (69). Change in the relative standardized uptake value (SUVmax; mean) for breast lesions imaged through neoadjuvant chemotherapy(69). Pathological response scored according to the Sataloff scale which categorises grades A complete response (blue), B good response (red) and nonresponse grade C + D (green). A significant difference was observed in breast uptake across the response groups ($p<10^{-4}$), noticeable from after the first cycle of chemotherapy ($p<10^{-5}$). B Figure adapted from Rousseau et al 2011 ((70). Change in the relative standardized uptake value (SUVmax; mean) for axillary lesions during neoadjuvant chemotherapy. Pathological response scored according to Sataloff scale; grade B good response (blue) and nonresponse grade C + D (green). A significant difference was observed in axillary nodal uptake across the response groups ($p<10^{-5}$), noticeable from after the first cycle of chemotherapy ($p<10^{-7}$)

EORTC PET guidance proposed thresholds for partial metabolic response dependent on interval between response assessment scan and first treatment, suggesting a ≥15% decrease in SUV after one cycle or ≥25% after more than one cycle to differentiate responders (74). PERCIST recommends uniform threshold change of ≥30% reduction in SULpeak that is not adjusted according to on treatment interval (62). Both sets of guidance apply generically across the spectrum of malignant disease but recognise need for tumour specific data. The published work in breast cancer reports a breadth of threshold SUV change for pCR prediction ranging from 15% to >80% of baseline at 1 cycle (Appendix 11.1). Phenotype specific data at this time point is lacking. Nevertheless the data overall suggests discriminatory SUV change may exceed EORTC and PERCIST thresholds (55). In practice relatively higher thresholds may be required for sub-selection for the best responders (low false -ve rate, 1-specificity, type 1 error) with a

view to therapy de-escalation compared to identification of non-responders with a view to offering escalation or transition to novel therapies within future clinical trials.

Breast and regional nodes are commonly involved in TNBC and an interpretation strategy for PET is required for multiple lesions. Most publications considered breast lesion only response but more recent studies have adopted an alternative strategy defining target as the single lesion with the greatest SUVmax at baseline, accepting that axilla rather than breast may be the index lesion (57, 75). PERCIST recommends evaluating up to five lesions but considers only % difference in SUL between the most intense tumour on study 1 and study 2 for response assessment. This lesion may differ between scans provided both lesions were present on the two studies and no progression is evident on the study 1 target.

Compared to dynamic analyses SUV assessment has the advantages of rapid scan duration but risks introducing errors due to potential therapy impact on tracer metabolism. A retrospective evaluation of dynamic FDG PET in 14 patients with mixed breast cancer who underwent 60 minutes dynamic acquisitions at baseline and midpoint of neoadjuvant chemotherapy has been performed (76). Scans were evaluated using a 2 compartment model, using a 1.5cm VOI placed over maximal tumour activity and blood input was derived from the cardiac blood pool. FDG transport (k1, ml/min/g) and FDG metabolism flux constant (K1, ml/min/g) were selected as response measures for comparison with DCE-MRI parameters and pathological response in the breast. Changes in tumour size, metabolism and vascularity measured by FDG PET and DCE MRI were well correlated and predictive of pathological response suggesting kinetic parameters may potentially provide further prognostic insight. In mixed breast phenotypes, model optimisation software was used to evaluate FDG Ki and compare with SUVpeak measured derived from a 45-60 minute acquisition performed at baseline and midpoint of neoadjuvant treatment (77). Average Ki and SUV were both greatest in ductal, high grade and hormone receptor negative tumours and changes in both dynamic and static parameters predicted subsequent pathological response, however only Ki change was able to predict later relapse.

3.2.3.2.3 Quality assurance in serial FDG PET acquisition and scan evaluation

European Association of Nuclear Medicine (EANM) guidance recommends common quality control/quality assurance procedures to ensure consistency in patient preparation, scan procedure and image reconstruction for FDG image interpretation (78). Conventionally FDG PET-CT scan acquisition commences 60-90 minutes after tracer injection, but the optimal timing in breast cancer is uncertain. Data derived from dynamic acquisitions in 40 patients with newly diagnosed locally advanced breast cancer demonstrated approximately linear rise

in SUVmax over the 27-75 minute acquisition time interval (79). However the rate of change differed between individual tumours such that in those with low initial uptake SUVs showed minimal change with acquisition time but up to 25% SUV increase over the final 15 minutes was observed in those with high early uptake. Consequently, failure to adhere to uniform imaging acquisition time across sequential scans may differentially confound interpretation of SUV change and a tolerance of +/-5 minutes for response assessment is recommended (78). Ongoing uptake has been reported several hours after tracer injection particularly in high grade breast tumours (80-82). Relationships between breast cancer receptor status and time-uptake are undefined but it is possible later image acquisition may optimise diagnostic accuracy at baseline and subsequent response assessment in TNBC.

In the majority of breast cancer response publications (Appendix 11.1), methodology used to define lesions suitable for PET analysis is either not stated or based on visual scan inspection by an experienced nuclear medicine reader. Humbert et al. (2014) (61) pre-defined hypermetabolic lesions suitable for follow up response imaging as those where baseline tumour SUVmax exceeded hepatic SUVmax. Alternative strategies require target lesions to exceed a particular size or SUV value (73, 83). PERCIST suggests baseline tumour uptake must exceed a threshold calculated from mean background, and background uptake within 20% of baseline on subsequent response scans (62). Specific application of PERCIST requirements has not been described in any of the publications listed in Appendix A.

A literature search identified no published test-retest data reporting intrinsic variability of FDG PET in breast cancer generally or its phenotype subsets. Meta-analysis of data from predominantly thoracic and GI sites report the combination of an absolute change in SUVmean of >1.2 units and relative change of >20%, and for SUVmax relative change of >30% and 2 unit absolute change exceeds the 95% test-retest variability and is likely to indicate a true therapy change rather than measurement error (84). In this meta-analysis SUVmean performed better than SUVmax, and factors including baseline high uptake and a delineation method using an isocontour technique were associated with improved repeatability and may be relevant to TNBC. However given the heterogeneity of uptake within breast cancer, TNBC phenotype specific data with regard to scan parameters and their repeatability is required to ascertain the biological significance of therapy induced change in SUV biomarker.

It is possible that SUV evaluation may be impacted by the chosen therapeutic independently of tumour response. The majority of published data (Appendix A) report FDG PET response following anthracycline first sequencing or mixed chemotherapy protocols. In the only chemotherapy comparative data, the interval PET was performed prior to the scheduled therapeutic transition in 60 women receiving taxane-anthracycline (Group A) or the reverse

sequence (Group B) (85). pCR rates did not differ between the two treatment groups. The mean change in SUVmax after 4 cycles of chemotherapy was 87.7% in those achieving pCR vs. 27% in those who did not ($p<0.01$). In contrast no significant difference in SUVmax response according to pathological response was present for women in arm A.. Limitations of this study include inconsistent scan acquisition time post tracer injection (median 60 minutes range 45 to 75 minutes) and disparity in ER and nodal status between groups (71% vs. 52% ER+ve and 90% vs. 10% node positive in Groups A vs. B respectively). Nevertheless these data suggest choice of therapeutic agent may itself influence SUV uptake and should considered when interpreting FDG PET response.

3.2.4 3'deoxy-3'-[18F]fluorothymidine (FLT) as a candidate imaging biomarker

FDG is not tumour specific and changes in uptake may be confounded by increased uptake due to post-therapy inflammation within the tumour or surrounding tissues, potentially falsely indicating a lack of response (44). Strategies with greater specificity for tumour metabolism may provide a better imaging surrogate for early tumour response in TNBC. The FLT tracer is an imaging biomarker of proliferation, reflecting use of thymidine (86). FLT enters cells by a combination of passive diffusion and active transport (ENT1), is phosphorylated in S-phase by thymidine kinase and trapped in its monophosphate form without incorporation into DNA. Recent meta-analysis has confirmed correlation between FLT uptake and the Ki-67 proliferation marker at a number of primary tumour sites including breast (87) but relationships with other cycle specific proliferation biomarkers have not been reported to date. Reduction in tumour Ki-67 during neoadjuvant chemotherapy is well recognised, and there is some evidence indicating that the magnitude of change is greatest in those who respond to treatment (37). However whilst Ki-67 change has been used as a pharmacodynamic marker in window of opportunity or neoadjuvant trials there are little validated data to define clinically meaningful change in response to cytotoxic therapy (88).

Review of published data identified nine studies in breast cancer. In unselected breast cancer a single study confirms FLT repeatability (SUV acquisition at 90 minutes, mean difference 10.5%, intra-class correlation coefficient 0.99) (89) and change in FLT uptake at one to two weeks following therapy initiation predates imaging or marker response in advanced disease (Table 1) (89-92). However ability of FLT tracer to predict pathological tumour response following neoadjuvant treatment is uncertain (93, 94) and phenotype specific data is required. Nevertheless as mitotic rate, which is typically high at baseline in TNBC, therapy impact on high baseline proliferation would be expected to occur rapidly following exposure to effective

systemic therapy suggesting FLT SUV change might provide a particularly useful response surrogate in this population.

3.3 Molecular imaging as a HER2 diagnostic

3.3.1 Rationale for HER2 targeted tracers for breast cancer imaging

The HER family of transmembrane tyrosine kinase receptors (HER1-4) are collectively important for regulation of cell survival, growth and differentiation. HER2 receptor kinase activation follows ligand-mediated heterodimerisation with another HER family member, or by complex formation with other transmembrane receptors particularly where the receptor is overexpressed or mutated. The transphosphorylated intracellular domains then interact with intracellular signalling molecules to activate or inhibit downstream pathways and cross talk with other signalling pathways leading to increased cell proliferation, cell motility, tumour invasiveness, progressive regional and distant metastases, accelerated angiogenesis and reduced apoptosis (95).

The humanised murine monoclonal immunoglobulin trastuzumab (Herceptin®, Roche) was the first HER2 targeted therapy to enter routine clinical practice. Following impressive clinical data in the first line metastatic setting, trastuzumab in combination with chemotherapy established a clear place in treatment of advanced HER2+ve disease (96, 97). Subsequently, combined hazard ratios (HR) from eight adjuvant/neoadjuvant studies involving 11,991 women demonstrated robust benefit for trastuzumab-containing regimens over chemotherapy alone (HR for OS and DFS 0.66; 95% CI 0.57 to 0.77; P<0.00001 and 0.60; 95% CI 0.50 to 0.71; p<0.00001 respectively) (98). Given the magnitude of benefit for trastuzumab, HER2 evaluation according to American Society of Clinical Oncologists/College of American Pathologists (ASCO/CAP) guideline recommendations is a mandatory diagnostic tissue assessment for all patients with newly diagnosed breast cancer (99).

Alternative approaches targeting the HER2 signalling pathway have been developed over the last decade with the aim of addressing the clinical challenge of metastatic progression due to trastuzumab or cytotoxic resistance. Of these agents lapatinib (Tyverb®, GlaxoSmithkline; oral HER1/HER2 receptor tyrosine kinase inhibitor), Pertuzumab (Perjeta®, Roche; a humanised monoclonal antibody binding at the subdomain 2 epitope of the HER2 extracellular domain) and the immunoconjugate trastuzumab emtansine (T-DM1; Kadcyla®, Roche) have demonstrated benefit in phase III trial evaluation (100-103). With this evolution of HER2

targeted therapies coordinated transition though approved agents at radiological evidence of progression has improved survival for those with metastatic disease (9). Despite this, patients will ultimately develop resistance to targeted approaches and metastatic HER2+ disease remains incurable. Potential predictors of trastuzumab resistance have been proposed, but no molecular marker has currently been validated in prospective clinical trials as a means for targeted therapy selection at metastatic progression. Consequently HER2 positivity defined according to ASCO/CAP guidance (99) remains the most important predictive factor for response to the spectrum of HER2 targeted therapies in routine clinical practice.

Practical challenges associated with tissue acquisition from metastatic sites means historical HER2 status assessed from the original primary breast tumour must frequently inform treatment decisions at relapse. Assumptions of constancy in gene expression with progression to metastatic disease have therapeutically significant implications, potentially increasing risk of futile use or inappropriate omission of HER2 targeted agents for some individuals (10, 11). Recent meta-analysis of 48 studies (2987 tumours) investigating the stability of HER2 expression between primary and metastatic disease reports pooled discordance proportion of 10% for distant metastases (95% CI: 7–14%), and 6% for loco-regional relapse (95% CI: 3–9%) ($p = 0.039$) (104). Use of both IHC and FISH testing, compared to IHC alone, did not improve concordance (discordance rate 10% vs. 5%, $p=0.02$) making it unlikely that technical errors explain this finding. Loss of HER2 expression was more common in studies using both IHC and FISH, but the pooled proportion of negative and positive HER2 conversion was 15% (95% CI: 10–21%) and 7% (95% CI: 5–10%) respectively indicating that initial HER2 status does not reliably predict direction of change in expression at relapse (104). Furthermore heterogeneous amplification, defined as the existence of two distinct or intermixed clones of breast cancer cells exhibiting different patterns of gene amplification is a well-recognised challenge to HER2 diagnostics in primary disease (105) and raises questions about HER2 heterogeneity across metastatic sites. Retrospective autopsy series suggest congruence across metastatic sites to be good but contain very few patients overall, and numbers with initially HER2+ primary disease are in single figures (106, 107).

Where metastatic tissue can be obtained sampling is invariably single site and single time point due to shared patient and clinician feasibility, safety and acceptability concerns precluding synchronous tissue acquisition from multiple sites within the same individual. Molecular imaging is a clinically attractive strategy for determining expression and localization of HER2-overexpressing tumour lesions, and potentially permits non-invasive evaluation of receptor status across multiple and difficult to biopsy sites at a time point relevant to targeted therapy selection. Despite phenotype associated differences in breast cancer uptake and some

evidence supporting ability to predict response following neoadjuvant HER2 targeted agents (83), FDG has no role for predicting HER2 status of primary disease or for discriminating HER2 status across metastatic sites. Consequently there is a clinical need for validated SPECT or PET radio-ligands targeting the HER2 phenotype that can inform targeted therapy selection and response evaluation. Ligands in preclinical or early phase clinical development include radiolabelled immunoglobulins (trastuzumab and pertuzumab), immunoglobulin fragments, and the non-immunoglobulin scaffolds, affibody and designed ankyrin-repeat proteins (108, 109). However no tracer has currently established its place in the clinic as a HER2 diagnostic outside the research context.

3.3.2 Antibody- based HER2 imaging

The high affinity and specificity for cell surface receptor targeting of intact monoclonal antibodies are ideal properties for molecular imaging. Trastuzumab is the most evaluated targeting ligand used for HER2 molecular imaging, in addition to its mainstream therapeutic role. SPECT or PET imaging using trastuzumab radiolabelled with ^{111}In (t½ 2.8 days) and ^{89}Zr (t½ 3.3 days) can differentiate high and low HER2 expression in subcutaneous xenograft models (110-112). HER2 downregulation in response to experimental heat shock protein 90 inhibition can be quantified using sequential PET (113) and, in trastuzumab treated MDA-MB-361 xenografts, sequential SPECT using ^{111}In labelled pertuzumab sensitively detected early molecular response at a time point which significantly predated later IHC confirmed decrease in viable HER2 positive cells (114).

The first clinical evaluations used ^{111}In-DTPA-trastuzumab to investigate the ability of SPECT for prediction of trastuzumab induced cardiac toxicity (115). Although unsuccessful in this regard, previously undiagnosed metastatic sites were detected in 13 of 15 baseline scans. Despite this only 45% of known tumour sites were evaluable but it was postulated that superior spatial resolution of PET might improve on the relatively low SPECT detection rate. In a feasibility study seeking to determine the optimal dosage and time of administration of ^{89}Zr labelled trastuzumab, most known tumour lesions in the liver, lung, bone and brain and previously unknown brain and bone lesions were visualised on PET performed at 5 days post tracer injection (116). This study did not incorporate tissue biopsy but the currently recruiting clinical trial NCT01420146 includes molecular characterization of tumour samples with discordant FDG and HER2 PET-CT image. Findings from this trial will provide specificity data for ^{89}Zr-trastuzumab in HER2 positive metastatic disease.

The high molecular weight (150 KDa) of monoclonal antibodies leads to relatively slow tissue penetration and long circulatory half-life giving the disadvantage of early images with very high

background signals. Preloading with unlabelled trastuzumab affords some improvement in bio-distribution (116, 117) but despite this antibody kinetics inherently necessitate use of isotopes with long half-lives that permit image acquisition at least several days following tracer administration to ensure optimal tumour-to-blood ratio. The shorter half-life of ^{64}Cu (t½ 12.7 hours) potentially permits a target scan time of only 48 hours post tracer injection. The first clinical data in 6 patients with at least 1 measurable non-contemporaneously tissue verified HER2+ breast cancer lesion confirmed feasibility for HER2 lesion evaluation despite relatively high liver uptake in trastuzumab naïve patients. (118). Further feasibility data for PET-CT using ^{64}Cu-DOTA-trastuzumab PET/CT at 48 hours in 8 patients with biopsy confirmed metastatic HER2+ve breast cancer reported detection sensitivity of 89% on day 2 post tracer injection, comparable to the detection sensitivity of 93% seen with FDG. Characterisation of tracer performance in a control HER2 negative population is planned (117).

3.3.2.1 Antibody Fragments

Imaging at 2 to 5 days post tracer injection poses significant feasibility and patient acceptability challenges that limit wider application of antibody imaging outside the clinical trial context. The critical antibody component for imaging application is the antigen binding sites located on the tip of the antigen binding (Fab) arms. Engineering fragments with lower molecular weight and shorter half-life through removal of the redundant Fc IgG domain may facilitate earlier imaging without compromise to antigen binding (Figure 2.1) meeting the clinical aspiration for same day HER2 imaging (119). As an example, F(ab)2 antibody fragments have reduced uptake in the normal organs and clear 4 to 5 times more rapidly from the circulation and tumour sites than the corresponding IgG form (120). In HER2+ xenografts, PET imaging at 24 hours after F(ab)2 trastuzumab fragments DOTA conjugated with ^{68}Ga (t½ 1.14 hours) was able to report HER2 downregulation in response to HSP inhibition that predated 18F-FDG PET glycolytic change (121). In subsequent clinical feasibility evaluation, only 7 lesions in 4 of 7 women with tissue confirmed HER2 positive metastatic lesions showed appreciable uptake at one, two and three hours following tracer injection (122). Sequential venous sampling confirmed relatively fast F(ab)2 clearance from the intravascular compartment and this combined with concomitant therapeutic trastuzumab for most of the HER2+ participants led the authors to postulate that competition for tumour targeting and rapid circulatory clearance may have resulted in inadequate time for extravasation and tumour localisation. To address this dose adjustment in the context of therapeutic trastuzumab and use of a radionuclide with longer half-life permitting imaging at 6-24 hours is planned.

Even smaller antibody variants may improve imaging properties for clinical application. Single-chain variable fragment (scFv) are fusion proteins comprising variable regions of the heavy (VH) and light chains (VL) of immunoglobulins connected with a short linker peptide of ten to about 25 amino acids. Diabodys, formed by linking two scFv, appear promising in their ability to function as an effective PET radiotracers. In pre-clinical murine models several agents have demonstrated HER2 targeting specificity and decreased PET signal in response to trastuzumab treatment (123-125) but have yet to be clinically evaluated. Nanobodies, the smallest (12-15KDa) naturally derived antigen-binding fragment, are characterised by high stability, rapid targeting and fast clearance. In the preclinical setting the PET tracer ^{68}Ga-NOTA-2Rs15dnti-HER2 delivered high specific contrast imaging of HER2+ tumours at 1 hour after tracer injection and progression to clinical validation is planned (126).

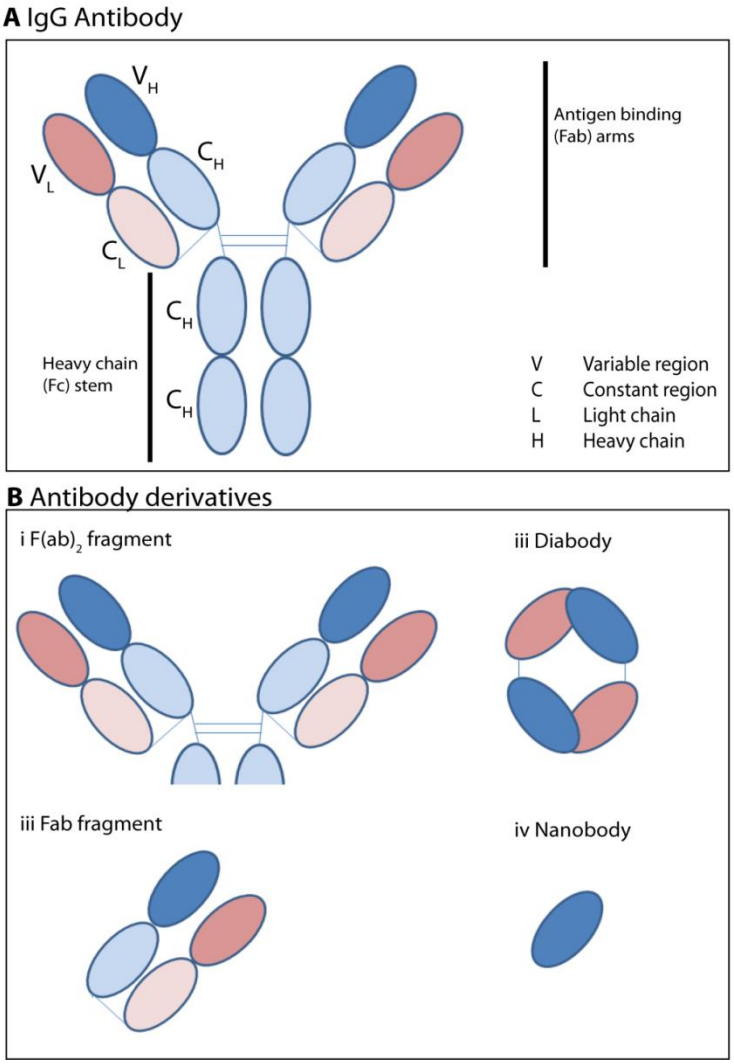


Figure 3-4 Schematic diagram of the IgG Antibody and its derivatives for use in molecular imaging

3.3.3 Non antibody derivatives

It is possible that imaging requirements of efficient tracer extravasation, good tissue penetration, and fast blood clearance may require molecules even smaller than the mass of the smallest immunoglobulin fragments.

3.3.3.1 Affibodies

Affibodies are small (6-7KDa) 58 amino acid molecules with a non-immunoglobulin scaffold originally derived from the 58 amino acid IgG binding domain of staphylococcal surface protein A. Their 13 amino acid binding surface can be manipulated to give high affinity binding to a variety of antigen targets and good imaging properties (108, 127). In the preclinical setting, HER2 targeting affibodies have demonstrated efficient pharmacokinetics with rapid extravasation to tumour targets and blood clearance permitting optimal image acquisition as early as 30 to 60 minutes following tracer injection and (128-130). In murine models several affibody PET and SPECT tracers have demonstrated ability to assess HER2 expression and monitor possible changes of receptor expression in response to therapeutic interventions providing better contrast in HER2 imaging than trastuzumab (128, 131-134).

The first clinical data using HER2 targeted affibody imaging was reported in 2010 (135). Three patients received ^{111}In- and/or ^{68}Ga-labeled DOTA-$Z_{HER2:342}$ (ABY-002) prior to SPECT or PET imaging. The first half-life was within 15 minutes of injection for both forms and rapid tracer kinetics permitted high contrast image acquisition 2-3 hours following injection. More recently, ^{111}In-ABY-025 underwent first-in-human study evaluation in 7 patients with FDG avid breast cancer metastases (136). SPECT acquisitions at 4, 24 and 48 hours following tracer injection were able to discriminate HER2 status of metastases and correctly predicted HER2 negative transition in one patient. However high hepatic as well as renal background uptake obscured sites of liver metastasis (135). Given the liver is commonly involved at HER2+ relapse this may limit clinical application of current affibody tracers and a current focus is developing constructs with reduced liver biodistribution (137).

3.3.3.2 Designed ankyrin repeat proteins (DARPins)

DARPins are small recombinant non-antibody based ankyrin repeat proteins with low molecular weight (14-18 kDa) that can be radiolabelled for imaging applications. The G3 DARPin has been developed as a novel HER2 imaging agent to subdomain 4 of HER2 ECD and has picomolar affinity for HER2 binding (138). In preclinical murine models excretion is predominantly renal and its short intravascular half-life permits same day imaging. Preclinical

evaluation in HER2+ ovarian cancer xenografts established using the SKOV3 cell line demonstrates high tumour:blood ratios at 1 hour and good tumour visualisation (139). Moreover promising in vivo biodistribution and HER2 discrimination has been reported for the ^{111}In DOTA conjugated $(HE)_3$-G3 DARPin (^{111}In DOTA- DARPin) following evaluation in BT474 (HER2+ breast cancer cell line) subcutaneous xenograft models and first-in-man clinical evaluation is planned (140, 141).

3.4 Rationale for work undertaken in this thesis

Molecular imaging offers the potential of optimisation and rationalisation of therapy at the individual patient level to facilitate therapy individualisation, reduce risk associated with complex metastatic site tissue acquisition by biopsy and ultimately reduce the of burden lengthy clinical trials to determine efficacy of novel agents. However tracer performance must be validated prior to clinical application as diagnostic and/or surrogate endpoint biomarkers.

There is limited clinical data reporting sensitivity and specificity of HER2 targeted tracers with reference to pathological confirmation of HER2 status at both visualised and non-visualised metastatic sites and specifically whether tracers can accurately reflect heterogeneity of overexpression across multiple metastatic sites within individual patients. Questions concerning histological verification in primary and regional nodal disease may be addressed using a study design that relates pre-surgical imaging to subsequent definitive histology from the primary tumour and regional nodes. However, in metastatic disease individual lesion accessibility and the safety and acceptability of tissue acquisition at multiple sites will significantly constrain the ability to acquire sensitivity and specificity data for these tracers within the context of a clinical study. The G3 DARPin is scheduled to enter clinical evaluation but despite promising preclinical data in a single HER2 positive cell line further supporting preclinical data would be advantageous prior to the first-in-man clinical trial. The results chapter's five present experiments undertaken to establish pre-clinical in vivo metastatic models to provide additional data relating HER2 tracer uptake to standard imaging techniques and histology in order to inform subsequent tracer application for clinical diagnosis.

In TNBC there is a clear clinical aspiration for early surrogate markers of the pCR and quantitative residual disease endpoint that can be applied to optimise neoadjuvant response assessment. Current evidence suggests FDG and FLT PET tracers may be informative after only one chemotherapy cycle. However work is required to ensure tracers meet characteristics required for clinical application as a predictive surrogate biomarker in TNBC before either can be adopted in routine clinical practice. Specifically breast cancer subtype specific experiments are required to provide data concerning tracer intrinsic variability, the optimal SUV reporting parameter and to define level of SUV change required for differentiating categories of later pathological response. Results chapters six to nine describe work undertaken to address these unanswered questions in the TNBC phenotype.

4 Pre-Clinical Experimental Material and Methods

4.1 Cell lines and cell culture

4.1.1 Cell lines

Five human breast carcinoma cell lines were selected for use in this study in order to represent a spectrum of HER2 expression levels. The breast cancer cell lines BT474, HCC1954, SKBR3 and HCC1143 were provided courtesy of the Breakthrough Breast Cancer laboratory at KCL (originally purchased from the American Type Culture Collection, ATCC, Teddington, UK). The breast cancer cell line MDA-MB-231 was a kind gift from Dr. Gilbert Frühwirth, KCL. Clinical and pathological features of these cell lines are summarized in Table 2.1

Cell line	HER2 copy number (142)
HCC1954	45.01
SKBR3	31.0
BT474	52
MDA-MB-231	0.9
HCC 1143	1.08

Figure 4-1 HER2 status of Breast Cancer Cell lines use in this study

4.1.2 Cell culture

The breast cancer cell lines HCC1954 and HCC1143 were cultured in RPMI 1640 (Gibco®, Life Technologies, Paisley, UK) supplemented with 10% heat-inactivated foetal bovine serum (FBS) (Gibco®, Life Technologies, Paisley, UK). BT474 and MDA-MB-231 cells were maintained in Dulbecco's modified Eagles medium (DMEM) (Sigma-Aldrich, Gillingham, UK) supplemented with 10% Fetal Bovine Serum (FBS). The SKBR3 cell line was cultured in McCoys media (Gibco®, Life Technologies, Paisley, UK) supplemented with 10% FBS. Cell lines were grown in the presence of penicillin and streptomycin (Sigma-Aldrich, Gillingham, UK).

Cryopreserved cells were recovered from liquid nitrogen by rapid thawing at 37˚C. The suspension was then mixed with 1mL culture medium and cells were isolated by centrifugation at 1200 RPM for 5 minutes, and then re-suspended in medium prior to transferring them into tissue culture flasks (Helena Biosciences, Gateshead, UK). Cells were grown in a humidified incubator at 37˚C and 5% CO_2. Upon reaching 80% confluence, cells were passaged by washing

once with sterile Dulbecco's Phosphate Saline Buffer (PBS) (Sigma-Aldrich, Gillingham, UK), then adding 2-4 mL (according to flask size) of 1X Trypsine-EDTA (Sigma-Aldrich, Gillingham, UK) and incubating the cells at 37˚C for 1 to 5 minutes, depending on the cell line. Detached cells were then isolated by centrifugation at 1200 RPM for 5 minutes, and re-suspended in complete medium.

Cell density was measured by mixing 10µL of the cell suspension to an equal volume of 0.4% Trypan Blue Solution (Gibco®, Life Technologies, Paisley, UK) then 10µL of the mixture were placed into a Countess Cell Counting Chamber Slide (Life Technologies, Paisley, UK). Cell density and viability were determined using the dye exclusion test by the Countess™ Automated Cell Counter (Life Technologies, Paisley, UK).

For long-term storage in liquid nitrogen, cell line stocks were prepared at low passage numbers. 1 or 2 x10^6 cells were re-suspended in 1mL complete media supplemented with 10% dimethyl sulfoxide (DMSO, Sigma-Aldrich, Gillingham, UK) and transferred into cryovials, which were cooled at -80◦C overnight and subsequently stored in liquid nitrogen.

Regular screens were performed using the MycoAlert Mycoplasma Detection Kit (Lonza, Basel, Switzerland) to ensure that all cell cultures remained mycoplasma free.

4.1.3 Antibiotic kill curve estimation

Breast cancer cell lines were plated at 40,000 cells per well, into 8 wells of a 12 well plate. At 24 hrs post plating blasticidin was added at concentration of 0, 1, 2.5, 5, 7.5, 10, 15 and 20 µg per well (2-10µg/ml well recommended in manufacturer's instructions). The percentage of cells surviving at 48, 72 and 120 hours was visually estimated and used to generate kill curves from which the dose of blasticidin required for selection following lentiviral transduction was determined (Table 2.2).

Puromycin dose use for selection was taken from kill curves previously performed in the Breakthrough Laboratory by Dr Dan Weekes (HCC1954) and Nirmesh Patel (SKBR3), per Table 4.2 below.

Cell line	Blasticidin	Puromycin
SKBR3	10µg/ml	1.5µg/ml
HCC1954	7.5µg/ml	3 µg /ml
HCC1143	2µg/ml	
BT474	7.5µg/ml	
MDA- MB-231	7.5µg/ml	

Figure 4-2 Antibiotic doses used for selection after lentiviral transduction

4.2 HER2 Protein Expression analysis

4.2.1 Sample collection

Cells plated for 24 hours at 150,000 cells per well (6 well plate) were washed with PBS and then lysed in 200µl Laemmli buffer (0.2M Tris HCL, 8% SDS, 40% Glycerol, 10% β-mercaptoethanol, 0.4% blue bromophenol). Following collection samples were incubated at 95˚C for 5 minutes then rapidly cooled on ice prior repeated aspiration using an insulin syringe. Samples were stored at -20˚C.

4.2.2 Antibodies

The primary antibodies used in this study are listed in Table 4.3. HRP-conjugated mouse and rabbit secondary antibodies (RPN2124, GE Healthcare, Amersham, UK) were used at 1:50,000 dilutions in 5% non-fat milk (Marvel) in PBS 0.05% Tween50. HRP-conjugated mouse and rabbit secondary antibodies (RPN2124, GE Healthcare, Amersham, UK) were used at 1:50,000 dilutions in 5% milk (Marvel).

Antibody	Company	Source	Dilution	Molecular Weight
HER2/ERB2 (#2242)	Cell signalling technology	Rabbit	1:1000	185KDa
β-actin	Sigma-Aldrich	Mouse	1:30000	42KDa

Figure 4-3 List of primary antibodies used for western blot analysis

4.2.3 Western Blotting

Samples were loaded onto 5-15% acrylamide gels (mini-PROTEAN®TGX™ Precast Gel, Bio-Rad, Hemel Hempstead, UK) and electro-transferred onto nitrocellulose membranes (Bio-Rad, Hemel Hempstead, UK). Membranes were stained with Ponceau S solution (Sigma-Aldrich, Gillingham, UK) to confirm successful transfer. Following wash with 1xTBST membranes were blocked with 5% milk (Marvel) prepared in 1xTBST and this was followed with incubation with the primary antibodies listed in Table 2 for 1 hour at room temperature or overnight at 4˚C. Membranes were then washed six times with 1xTBST with and incubated in secondary antibodies diluted in 5% milk for 1 hour at room temperature. Following 6 further membrane washes with 1x TBST membranes were incubated in ECL prime (GE Healthcare, Amersham, UK) for visualisation. Hyper-films ECL (GE Healthcare, Amersham, UK) were exposed for varying times at the SRX-101A developer (Konica Minolta, Milton Keynes, UK).

4.3 Luciferase expression

4.3.1 Firefly Luciferase Lentiviral transduction

Breast Cancer Cell lines were plated at 150,000 cells per well in tissue culture 6 well plates (Greiner bio-one, Stonehouse, UK) and incubated overnight at 37˚C and 5% CO2. LVP 439 (EF1a-Luciferase [firefly]-2A-RFP [Bsd]) premade Lentiviral expression particles (Amsbio, Oxford, Uk) were thawed on ice and cells were then infected using a 1:3 virus:plain media ratio. The plate was incubated overnight at 37˚C and 5% CO2 and the infection medium was then replaced with fresh complete media the following morning. Blasticidin selection for transduced cells (according to the predetermined antibiotic kill curve) was performed after a further 24 to 48 hours. The selected cells were not used for experiments until a minimum of 3 days blasticidin selection.

4.3.2 *In vitro* confirmation of luciferase lentiviral transduction

Transduction was checked by visual estimation of RFP presence under a fluorescence microscope and confirmed by IVIS evaluation of the transduced cells lines (4.5.2)

4.4 *In vivo* breast cancer xenograft model establishment

Animal studies were carried out in accordance with UK Research Councils' and Medical Research Charities' guidelines on Responsibility in the Use of Animals in Bioscience Research, under UK Home Office license (PPL 70/7775; Title: Models of breast cancer heterogeneity and

biomarkers). Animals were maintained in either the Hodgkin or Rayne BSU, Kings College London.

Female six to eight week old CD-1 Foxn1-/- Nude mice were purchased from Charles River Laboratories (UK). The animals were maintained under sterile conditions in filter topped cages on sterile bedding and fed an irradiated diet of standard mouse chow.

Mycoplasma negative cell lines of interest engineered to stably express luciferase were used to establish *in vivo* tumour models for longitudinal imaging (bioluminescence or SPECT). To establish subcutaneous models, one million cells suspended in 50µl sterile PBS were injected into the mammary fat pad at day 0. To establish lung metastasis models mice were restrained and injected in the tail vein with 1 million cells suspected in 100µl PBS at day 0. Mice were monitored post procedure, the day following injection and at least once or twice weekly thereafter. Subcutaneous tumours were measured using calliper's (mm) and volume calculated using the formula Volume= (length*width)$^{2)}$/2). Mouse monitoring comprised assessment of weight, behaviour and palpable tumour size up to three times per week. In accordance with the PPL evidence of weight loss greater than 10%, moderate signs of behavioural distress and/or palpable tumour diameter ≥15mm (single tumour) or ≥10mm (multiple tumour) precipitated humane killing by Schedule 1 method.

4.5 Bioluminescence imaging

Bioluminescent imaging (BLI) was performed using (IVIS® Lumina Series III, Perkin Elmer). Imaging and quantification of signals was controlled by the acquisition and analysis software Living Image®.

4.5.1 Luciferin stock preparation

30mg/ml stock solution D-Luciferin Firefly, potassium salt (Calliper Life Sciences, Hopkinton, USA) for use in *in vitro* and *in vivo* experiments was prepared by dissolving 1g of D-Luciferin in 33.3mls DPBS without magnesium or calcium and mixed prior to filtering through a 0.2µm syringe filter (Merck Milliporem Watford, UK). Aliquots were frozen and stored at -80˚C. Prior to use aliquots were protected from light and thawed on ice.

4.5.2 *In vitro* imaging

Transduced bioluminescent cells lines of interest were trypsinised and counted to achieve stock dilutions of 1 million cells per ml. Cells were plated into black clear bottom 96 well

plates at decreasing dilutions from 10000 to 100 cells per well in 200µl of the appropriate complete media.

The next day 2 µl of D-Luciferin stock was added to each well (1:100 dilution) and the plate transferred IVIS™ Lumina Series III (Perkin Elmer) for BLI. Image times were 10-30 seconds per plate. For each cell line the light emitted per cell quantified from the ratio of photons per well and the total cell number.

4.5.3 *In vivo* bioluminescence imaging

Following cell injection at day 0, animals were serially imaged to monitor tumour formation using BLI (IVIS™ Lumina Series III, Perkin Elmer) until an experimental end point defined by severity limit (Tumour Max Diameter 10mm). Animals were imaged in cohorts of 3 to 5 mice. At each imaging session mice were intraperitoneal injected with D-Luciferin Firefly, potassium salt (Caliper life Sciences, Hopkinton, USA) 150mg/kg (30 mg/mL stock) at time 0. Animals were then anaesthetised in a clear Plexiglas anaesthesia box (2% isofluorane) and once fully anaesthetised transferred to the nose cones (2% isofluorane) attached to the manifold on the warmed stage of the light tight imaging chamber. Imaging times ranged from 1-30 seconds depending on tumour model and was performed at 6-12 minutes post luciferin injection. On completion mice were returned to their cages and monitored until recovery.

BLI was repeated weekly until the severity limit determined experimental endpoint, when mice were culled immediately following their final imaging session using a Schedule 1 method. Removal of all tissues of interest was confirmed by *ex vivo* BLI of tissues and the residual cadaver using the IVIS® camera system. Excised tissues of interest were fixed in 10% formalin and prepared for histopathology verification.

Light emitted from bioluminescent tumours or cells was detected by the IVIS® camera system, integrated, digitalised and displayed. Image analysis was performed using the Living Image® software (Xenogen). Serial scans across were reviewed and the colour scale normalised across the longitudinal series to facilitate visual comparison across scans. For quantitative comparison an appropriately sized region of interest (ROI) was selected to encompass subcutaneous tumour or lung areas and the total flux (photons per second) within these ROIs were measured.

4.6 Single Photon Emission Computed Tomography (SPECT) imaging

4.6.1 *In vitro* SPECT imaging

The NanoSPECT/CT® detection limit for cell lines imaged HER2 targeted radiotracer was determined. Tracer used in these experiments was either ^{111}Indium DOTA conjugated $(HE)_3$-G3 DARPin (^{111}In-DOTA-DARPin) or ^{111}In-CHX-A"-DTPA-trastuzumab (^{111}Indium trastuzumab) prepared by Dr Margaret Cooper (Appendix 4). Cell lines used in these experiments were HCC1954 and MDA-MB-231 cells engineered to stably express RFP and Firefly luciferase and maintained under blasticidin selection or HCC1954 engineered to stably express RFP and Firefly luciferase with HER2 knockdown maintained under puromycin selection.

Cell lines of interest were trypsinised and counted, 35 million cells were then re-suspended in 7ml of filtered PBS + 0.5% Bovine Serum albumin buffer (Fisher Scientific, USA), to achieve a stock dilution of 5 million cells per ml. The stock was used to prepare serial dilutions to attain triplicate suspensions of progressively decreasing cell number from 5 million to 5000 cells in 2ml Eppendorf tubes. ^{111}In-labelled tracer was added to each tube to achieve a 10 nanomolar tracer incubation concentration, 1500ul total volume in each tube. Triplicate reference 10 nanomolar standards were prepared at the same time and reserved.

To mimic *in vivo* conditions suspensions were then incubated for 30 minutes at 37˚C prior to centrifugation at 1200 RPM for 5 minutes. Pellets were washed twice in buffer and transferred to 500µl PCR tubes prior to a final spin and aspiration of the supernatant from each tube. Each tube was then positioned in a rack on the mouse bed of the NanoSPECT/CT Silver Upgrade (Mediso Ltd., Budapest, Hungary) and images acquired over a scan time permitting 60-70 seconds per frame.

On completion of their final scan, activity in tubes including reference standard was counted (Wallac 1282 Compugamma Perkin Elmer, UK). For each cell pellet size % tracer binding was calculated using (mean counts of the pellet triplicate as a percentage of mean counts of the triplicate reference standard). Following scan reconstruction SPECT image analysis was performed using VivoQuant™ version 1.23

4.6.2 *In vivo* SPECT imaging

Animals bearing subcutaneous mammary tumours or lung metastases were imaged using HER2 targeted radiotracer at a time point estimated from the IVIS data and animal monitoring to be optimal for tumour establishment. On the day of imaging for all experiments tracer was

prepared by Dr Margaret Cooper (see Appendix 11.5) and comprised either ^{111}In-DOTA-DARPin (G3 DARPin provided by Dr Robert Goldstein, UCL) or ^{111}In-trastuzumab.

On day of scanning, mice were transferred from the Rayne BSU to the imaging laboratory and anaesthetised with isofluorane, O2 flow rate of 1.0-1.5L/min and isofluorane levels of 2-2.5%. Under anaesthesia mice were injected intravenously with tracer by a laboratory worker wearing double gloves. Syringe weights prior to and subsequent to injection were recorded for each mouse to permit calculation of the delivered injected volume. Residual activity in the syringe and on gauze used to compress the injection site was measured using the capintec and recorded for each mouse.

For image acquisition, mice were positioned on the bed of the NanoSPECT/CT Silver Upgrade (Mediso Ltd., Budapest, Hungary) equipped with a multiplexed multipinhole (nine pinholes, aperture 1.0 mm) collimator. Imaging was performed under 2% isofluorane anaesthesia. Helical SPECT images were acquired over 30-40 minutes (SPECT 60-70 seconds per frame, CT images 55kVp, 1000ms and 1 degree angular stepping.). On completion of their final scan, mice were culled using a Schedule 1 method. The weight and activity of the culled mouse was measured. Mouse organs (gut, liver, kidneys, spleen, lungs, heart, tail and palpable tumour sites) were removed and weighed. The residual carcass weight and activity were noted. Tissues from anticipated tumour bearing sites (mammary gland or lung) were placed into formalin for subsequent histological verification and stored in the hot lab fridge; the remaining organs and the gauze used tracer injection were stored in the histology freezer.

Radionuclide images were reconstructed and fused with CT images prior to image analysis, performed using VivoQuant™ version 1.23 software.

4.6.3 Biodistribution calculation

Activity of the reserved organs, standards, gauze and background was performed using a gamma counter (Wallac 1282 Compugamma Perkin Elmer, UK) and used to calculate organ bio-distribution according to the method below

- The counts from the eight standards were plotted against the volume of tracer and the gradient of the line calculated. From this the tracer volume for each organ and the gauze was calculated using the formula volume=((organ count - background count)*gradient).

- The total injection volume determined by the difference in syringe weight pre- and post-tracer injection. The mouse injection volume was then calculated by subtracting the volume of tracer for gauze and tissue from the total injection volume.
- The formula ((organ volume/mouse injection volume)*100) was used to calculate the % injection volume for each organ
- % injected dose per gram was calculated by dividing the % injection volume by the measured organ weight.

4.6.3.1 Histology

Excised tissues of interest were fixed in 10% formalin. Preparation for histopathology examination (paraffin embedded, 3-4 micron sectioning, and staining) was performed by the Breakthrough laboratory technician, Erika Francesch Domenech. Stains used were haematoxyin (H&E) and eosin for metastases evaluation, Bond Oracle Leica Kit #TA9145 for HER2 evaluation and PECAM CD31 (ab28365) for vascularity.

H&E and HER2 immunostained slides representing a whole lung cross-section were scanned to the Digital Image Hub (Leica Biosystems, whole slide scanning system) for quantification of tumour burden. The scanned image was examined (JG) to identify all metastatic sites within lung tissues. Metastatic burden relative to the whole lung area was quantified by delineation of all tumour regions of interest (Figure 4.4) and the whole lung area using the polygon area evaluation tool. The sum of tumour areas was expressed as a percentage of the total lung parenchyma.

For vascularity scoring the most intense vascular areas (hotspots) on the CD31 stained tumour sections were selected subjectively from each CD31 stained tumour section and vessel count (Mean and SEM) quantified by counting the number of vessels present within three 0.3mm^2 grids placed within these areas.

A Delineation of lung metastases for area evaluation

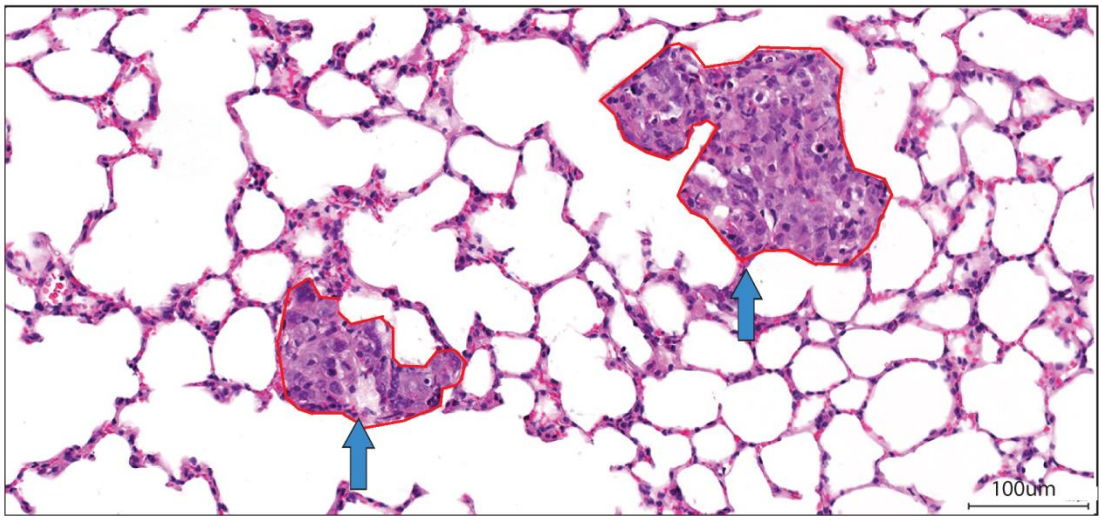


Figure 4-4 Quantification of lung metastatic burden

A. Lung H&E cross-section. The polygon drawing tool was used to delineate regions of interest around 2 metastatic sites (arrowed) for area evaluation using the Digital Image Hub (Leica Biosystems, whole slide scanning system).

4.7 Statistical analysis

Microsoft Excel was used to calculate mean, standard deviation and standard error of the mean and generate graphs for visual comparison of mouse groups. Other statistical analysis was performed using IBM SPSS version 22 statistical software. Normality was confirmed using the Shapiro-Wilk test. For parametric data means of two independent groups were compared using the independent samples t-test. Comparison of means between more than two independent groups was performed using one-way ANOVA (parametric data) and Tukey post-hoc tests.

5 Results: Preclinical HER2 imaging

5.1 Aims

The objective of the experiments described in this chapter was to characterise the ability of the HER2 targeted molecular imaging tracer to differentiate the spectrum of HER2 expression seen in different breast cancer phenotypes.

Breast cancer cells lines were engineered to express luciferase and assessed for their tumourogenicity in subcutaneous and intravenous (lung) metastasis models using longitudinal bioluminescent imaging (BLI) and *ex vivo* histological verification. The successful metastatic xenograft models generated by subcutaneous or tail-vein injection of HCC1954 and MDA-MB-231 breast cancer cell lines with constitutive lentiviral delivered luciferase expression were chosen to evaluate ^{111}In-DOTA conjugated $(HE)_3$-G3 DARPin (^{111}In DOTA-DARPin) SPECT radiotracer tracer performance (141). The experimental models were assessed by *in vivo* SPECT CT imaging and *ex vivo* radiotracer bio-distribution. Cell lines and *in vivo* tumour models were also evaluated using radiolabelled trastuzumab imaging control and *in vitro* tracer binding experiments were performed with cell lines of interest to facilitate interpretation of the *in vivo* results.

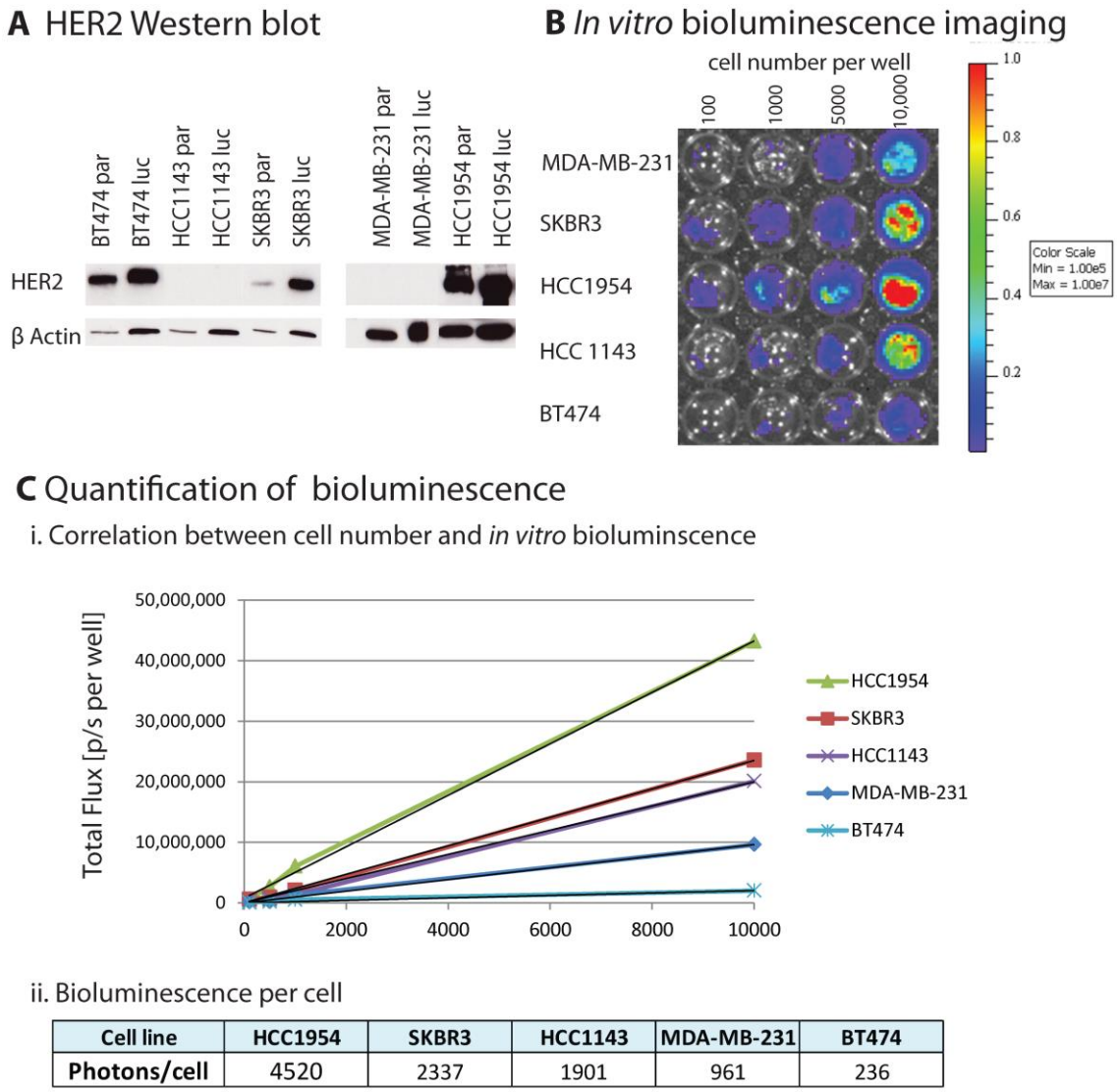
5.2 Development of bioluminescent metastatic models

5.2.1 *In vitro* evaluation of luciferase-transduced cell lines

The five breast cancer cell lines were infected with the firefly luciferase gene (LVP virus, Amsbio) and cells stably-expressing luciferase were selected as described in the methods section. HER2 status was confirmed with Western Blot of parental and Luc+ lines (Figure 5-1A). Luc+ cells from each line were serially diluted, plated and imaged as described Chapter 4. Bioluminescence images (Figure 5-1 B) demonstrate that at least 5,000 cells are detectable in all transduced cells lines, and as few as 100 cells are detectable in the Luc+ lines SKBR3 and HCC1954.

Light emission was quantified using the Living Image software. Good correlation between total flux (photons per second) and the total number of cells was demonstrated in all transduced cell lines (Figure 5-1C i). Light emission per cell was calculated from the ratio of detected photons per well to total cell number (Figure 5-1Cii). Only the Luc+ BT474 cell line exhibited expression below 500 photons/cell/s and expression was 4 to 19 fold higher for HCC1954,

SKBR3, HCC1143 and MDA-MB-231 than for BT474 reflecting the ability to visualise fewer cells *in vitro*.



ii. Bioluminescence per cell

Cell line	HCC1954	SKBR3	HCC1143	MDA-MB-231	BT474
Photons/cell	4520	2337	1901	961	236

Figure 5-1 *In vitro* bioluminescence of breast cancer cell lines expressing luciferase.

A. Western blots for the 5 selected cell lines confirming HER2 expression in parental (par) and luciciferase transduced (luc) BT474, SKBR3 and HCC1954 cell lines only.

B. SKBR3, HCC1954, MDA-MB-231, HCC1143 and BT474 Luc+ cell lines were plated 10,000 to 100 cells and were imaged for 10 seconds at 10 minutes after addition of luciferin to the media. Wells containing media only served as negative controls.

C. (i) Correlation between cell number per well and bioluminescence (photons per second per well (p/s per well)) for the five cell lines. R^2 =0.95 for BT474 and 0.99 for all other cell lines (ii) Calculated photons per cell per second for each cell line.

In vivo bioluminescent imaging to monitor subcutaneous tumour growth

Subcutaneous tumours were established in the HCC1954 and MDA-MB-231 groups (total of 4 tumours and 3 tumours per group respectively) and monitored by BLI and bi-dimensional calliper measurements as described (4.4). By day 14, bioluminescence signalled successful tumour establishment, which predated calliper measurements of tumour volume (figure 5-2 A and B). At day 56 the SKBRR3, HCC1143 and BT474 cell lines failed to establish subcutaneous tumours monitored by both palpation or bioluminescence and mice were culled.

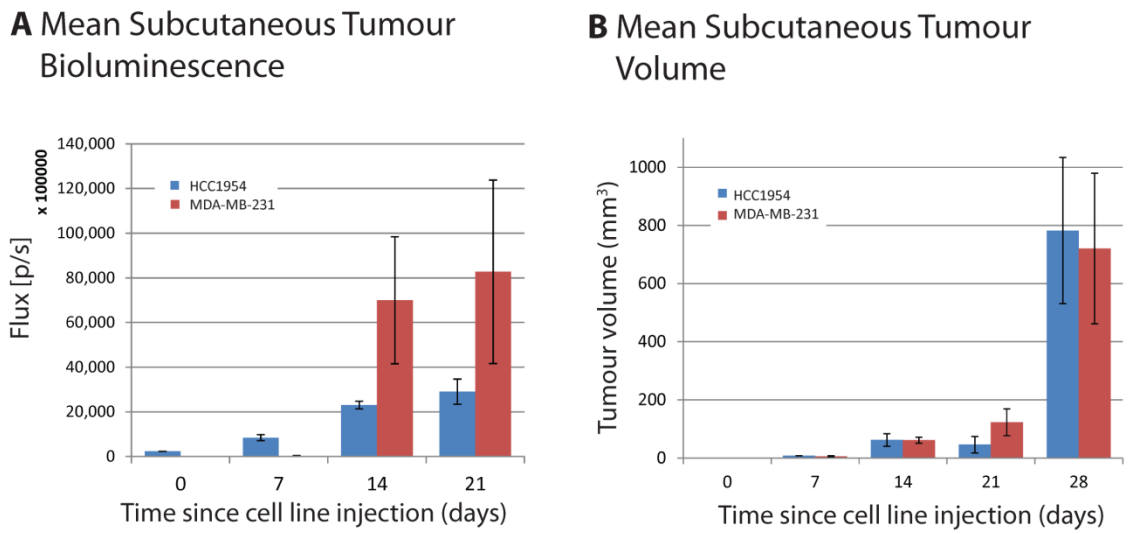


Figure 5-2 Monitoring subcutaneous tumour growth and bioluminescence *in vivo*

Data from mice bearing subcutaneous tumours established from Luc+ HCC1954 and MDA-MB-231 cells. Tumour growth was monitored by weekly bioluminescent imaging and calliper measurements (n=3 (HCC1954) or 4 (MDA-MB-231) tumours).

A. Sequential bioluminescence. Mean flux (photons per second) and SEM for each group are shown. Tumours were evaluable using bioluminescence by day 14 and further increase in total flux over time was observed in both groups. Technical failure on day 28 meant that final bioluminescence could not be performed prior to mouse cull (severity endpoint).

*B. Tumour volume was calculated from bi-dimensional calliper measurements (mm) using the formula Volume=((length*width)^2)/2). Mean volume (mm^3) and SEM for each cell line group are shown.*

5.2.2 *In vivo* bioluminescent imaging to monitor growth of metastases

The *in vivo* ability of Luc+ cell lines to establish lung metastases was evaluated using the HER2+ (HCC1954, SKBR3 and BT474) and HER2-negative (MDA-MB-231 and HCC1143) cell lines. One million Luc+ cells were intravenous injected into CD-1 Foxn1-/- Nude mice, 3 mice per cell line group on day 0. BLI was performed on day 0 and weekly thereafter until an experimental endpoint determined by severity limit.

Metastases were successfully established in all mice in the Luc+ HCC1954 and MDA-MB-231 cohorts (Figure 5.4). The day 0 images confirmed cell arrest in the lungs. Although the *in vitro* cell flux was five times greater in the HCC1954 cell line compared to the MDA-MB-231 cell line, the day 0 lung total flux did not demonstrate the same magnitude of difference. Consistent with the images, the measured flux in both groups initially rapidly declines from day 0, as cells are cleared from the lung and increasing thereafter as metastases are established from those cells that stay resident in the lung. This is most marked in the MDA-MB-231, potentially reflecting the faster basal growth rate observed in tissue culture for this cell line compared to the HCC1954 cell (MDA-MB-231 doubling time approximately three times that of HCC1954, 1.34 vs. 3.6 days). In both groups concerns over mouse weight and health necessitated organ harvest by day 56.

In the SKBRR3, HCC1143 or BT474 groups the day 0 images confirmed cell arrest in the lungs, however there was no subsequent BLI signal that mice had established metastases and mice were culled at day 82.

5.2.3 Histological confirmation of lung metastases

Metastases were confirmed *ex vivo* in H&E stained lung cross sections in all mice from HCC1954 and MDA-MD-231 groups (Figure 4.5). MDA-MB-231 tumour bearing-hosts had established metastases occupying mean 7% (SEM 0.02) lung cross sectional area, 0.74 metastases present per mm^2 lung parenchyma. In the HCC1954 group metastases covered 3% (SEM 0.01) of the lung cross sectional area, 0.24 metastases per mm^2 lung parenchyma. This confirms the greater efficiency of *in vivo* lung colonisation for the MDA-MB-231 cell line than HCC1954 that was suggested by the sequential *in vivo* BLI.

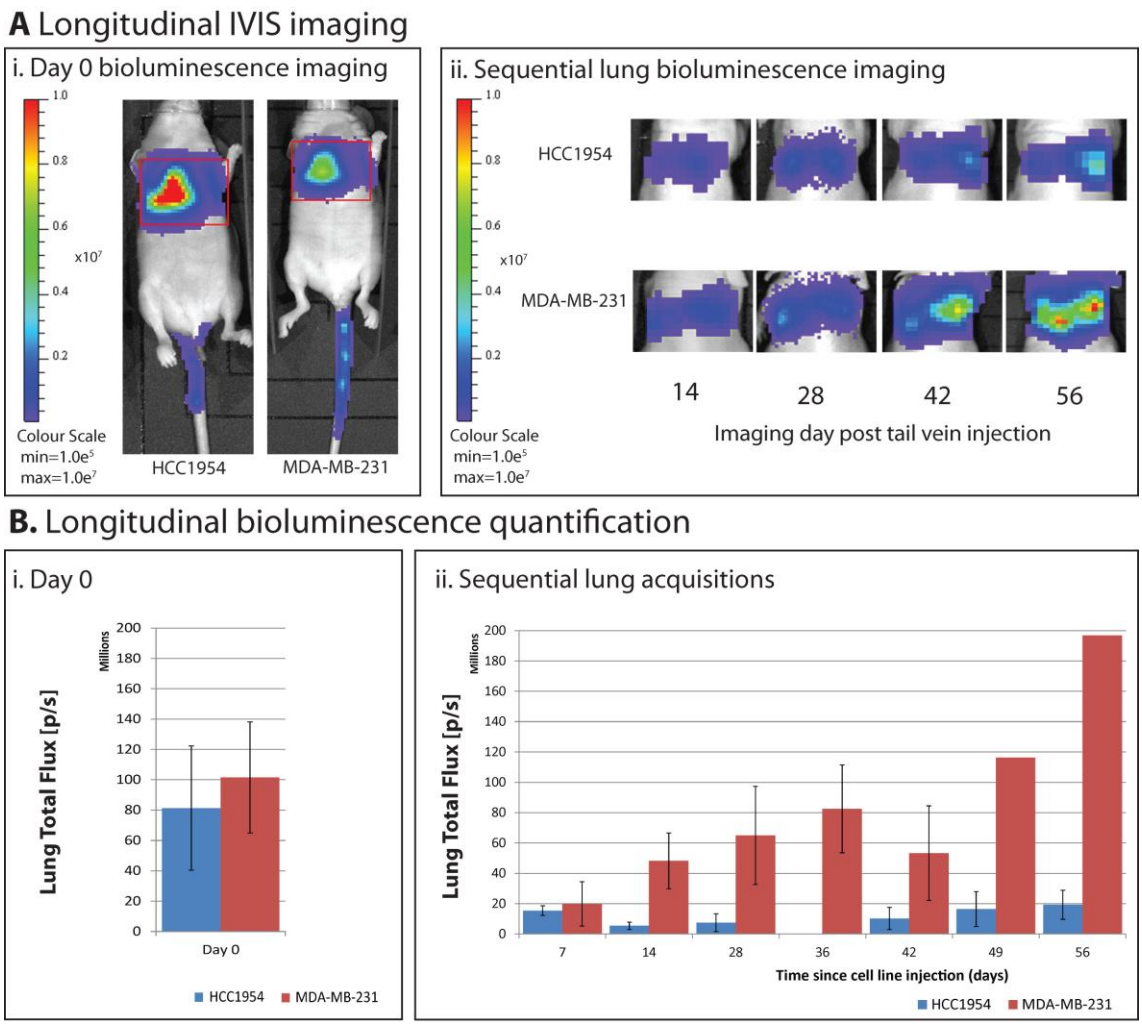


Figure 5-3 Longitudinal bioluminescence monitoring of lung metastases established following intravenous HCC1954 and MDA-MB-231 Luc+ cell line injection

A. (i) Representative whole mouse BLI on day of intravenous injection with the HCC1954 and MDA-MB-231 cell lines (day 0). Lung regions of interest (ROI) are delineated by red box. The day 0 image provides a visual confirmation that injected cells arrest in the lungs. (ii) Bioluminescent images showing lung areas from two mice intravenously injected with HCC1954 and MDA-MB-231 cell lines respectively. Bioluminescence declines from day 0 injection to a minimum at day 7-14 as cells are cleared from the lung that have not seeded as metastases. Bioluminescence incrementally increases thereafter as tumours are established.

B. Lung ROI were designated for each bioluminescent image over the experimental time course and quantified as total flux (photons per second, p/s) using Living Image® software (Xenogen). (i) Total flux (mean and SEM for mice scanned at day 0 following tail vein injection. (ii) Mice were scanned from day 7 onwards and quantification of these images used the same ROI size across all mice at all-time points. The mean and SEM for mice in each of HCC1954 and MDA-MB-231 groups are shown (n=3 per group). Consistent with the images, flux rapidly declines from day 0 in both groups, increasing thereafter as metastases are established.

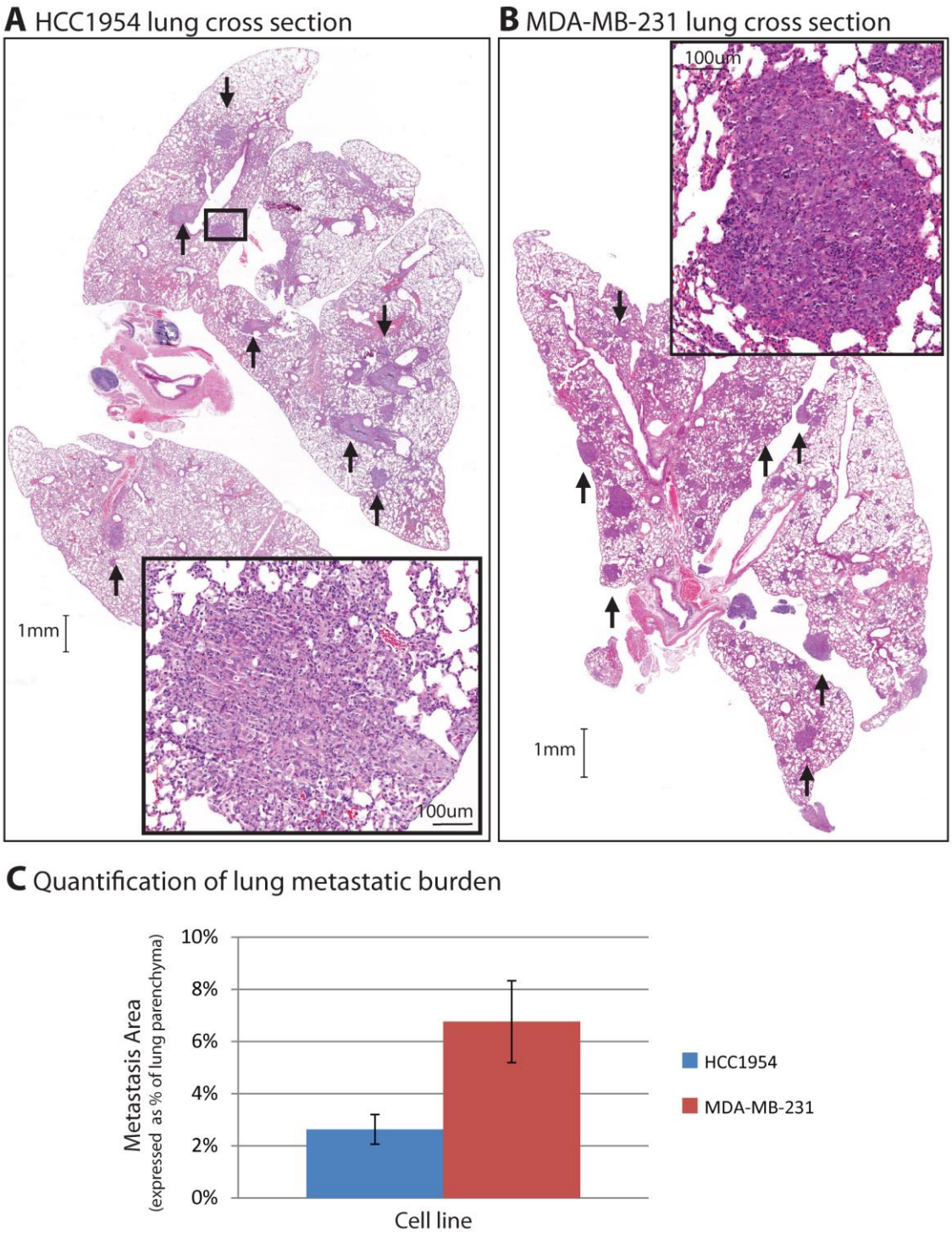


Figure 5-4 Histological verification of lung metastasis in HCC1954 and MDA-MB-231 models

Histological evaluation of the lung sections confirmed presence of metastases in all HCC1954 and MDA-MB-231 injected mice. Representative images (lung cross section (3µm) are shown. A. Low power view demonstrating multiple metastases (arrows) within the lung H&E cross section, HCC1954 model. Inset shows magnification of a single metastatic area within the cross-section (box). B. Low power view demonstrating multiple metastases (arrows) within the lung H&E cross section, MDA-MB-231 model. Inset shows magnification of a single metastatic area within the cross-section.

C. Metastatic area expressed as percentage of lung parenchyma area. Mean and SEM for HCC1954 and MDA-MB-231 cell lines.

5.3 HER2 molecular imaging of in *vivo* models

Based on the preceding bioluminescence and histology data, the HER2+ cell line, HCC1954, and the HER2-negative cell line, MDA-MB-231, with constitutive lentiviral delivered luciferase expression were selected for use in these experiments. Cell lines were confirmed to be mycoplasma negative and maintained under blasticidin selection prior to tail vein injection. The xenograft models were maintained in the Rayne BSU but decommissioning of the IVIS® machine on that site precluded BLI confirmation of metastasis formation prior to SPECT-CT using the HER2 targeting ^{111}In DOTA- DARPin for these cohorts.

5.3.1 Cohort 1 model preparation

An initial cohort of eight mice were taken through to SPECT-CT imaging and bio-distribution, 4 following tail vein injection of Luc+ HCC1954 cell line, 3 following tail vein injection of Luc+MDA-MB-231 cell line and 1 naïve control. Day 0 injections took place with a view to scanning at day 33 post-injection. This time point was selected based on the preceding IVIS cohort and it was intended to balance achieving reasonable disease burden whilst minimising respiratory compromise that could compromise anaesthesia for the duration of SPECT-CT imaging.

Day 0 tail vein injection of CD-1 Foxn1-/- Nude mice with Luc+ MDA-MB-231 cell line took place without incident; 3 injected mice progressing to scheduled imaging at day 33 (M1, M2 and M3). Two HCC1954 injected mice became immediately unwell following cell injection and third mouse experienced respiratory arrest within minutes of injection. To ensure adequate models for the scheduled imaging session three further mice were injected 3 days later by a different operator. Unfortunately injections were again poorly tolerated resulting in a further respiratory arrest and mouse death. To address this Fresh HCC1954 cell stocks were recovered from liquid nitrogen and passaged three times prior to tail vein injection in a further two mice, 2 million cells in 100µl in view of the 20 day interval to scheduled imaging without incident. In total 4 mice who received HCC1954 tail vein injections were imaged at day 33 (H1), day 30 (H2) and at day 20 (H3 and H4, third injection group).

5.3.2 SPECT-CT imaging

The mouse cohort was imaged over two days. ^{111}In DOTA-DARPin tracer was freshly prepared for injection on both days by an experienced radiochemist (Appendix 4). Imaging was performed as described in the methods and took place without incident in seven of the eight mice. Mouse M2 died approximately 30 minutes post tracer injection. This may have been due to a greater disease burden. To ensure a delayed HER2 imaging signal was not missed, one mouse (H3) was allowed to recover and a later acquisition was performed the following day, at 19 hours. Organs were harvested for tracer bio-distribution on completion of imaging in accordance with the methods.

5.3.2.1 Establishing a suitable time point for scan acquisition

With the aim of establishing a suitable time point for animal imaging following tracer injection, serial SPECT acquisitions were performed in two mice (H1 and H2) successfully injected with the HCC1954 cell line. H1 underwent serial imaging with 6 acquisitions over 3 hours prior to cull and organ harvest for bio-distribution (60 seconds per frame for each SPECT acquisition, total 30 minutes per scan) commencing immediately following tracer injection. However inspection of the thoracic blood pool (heart and descending aorta) suggested very rapid elimination of tracer from the circulation (Figure 4-1). Metastasis could not be visualised within the lungs at any time point on the serial images

A second mouse, H2, underwent two early acquisitions at 0-30 and 30-60 minutes and was allowed to recover prior to delayed acquisition at 19 hours following tracer injection (SPECT scans 60 seconds per frame, total 30 minutes per scan). Rapid tracer clearance from the blood pool at the early 30-60 minute time point was confirmed. Based on these data, the 30-60 minute time point was determined to be a practical time point for SPECT acquisition and was adopted with all subsequent mice. However, tumour could not be visualised within the lungs at any time point on the acquired images

5.3.2.2 SPECT metastases evaluation

Presence of metastasis could not be determined from the image acquisitions in any mouse model.

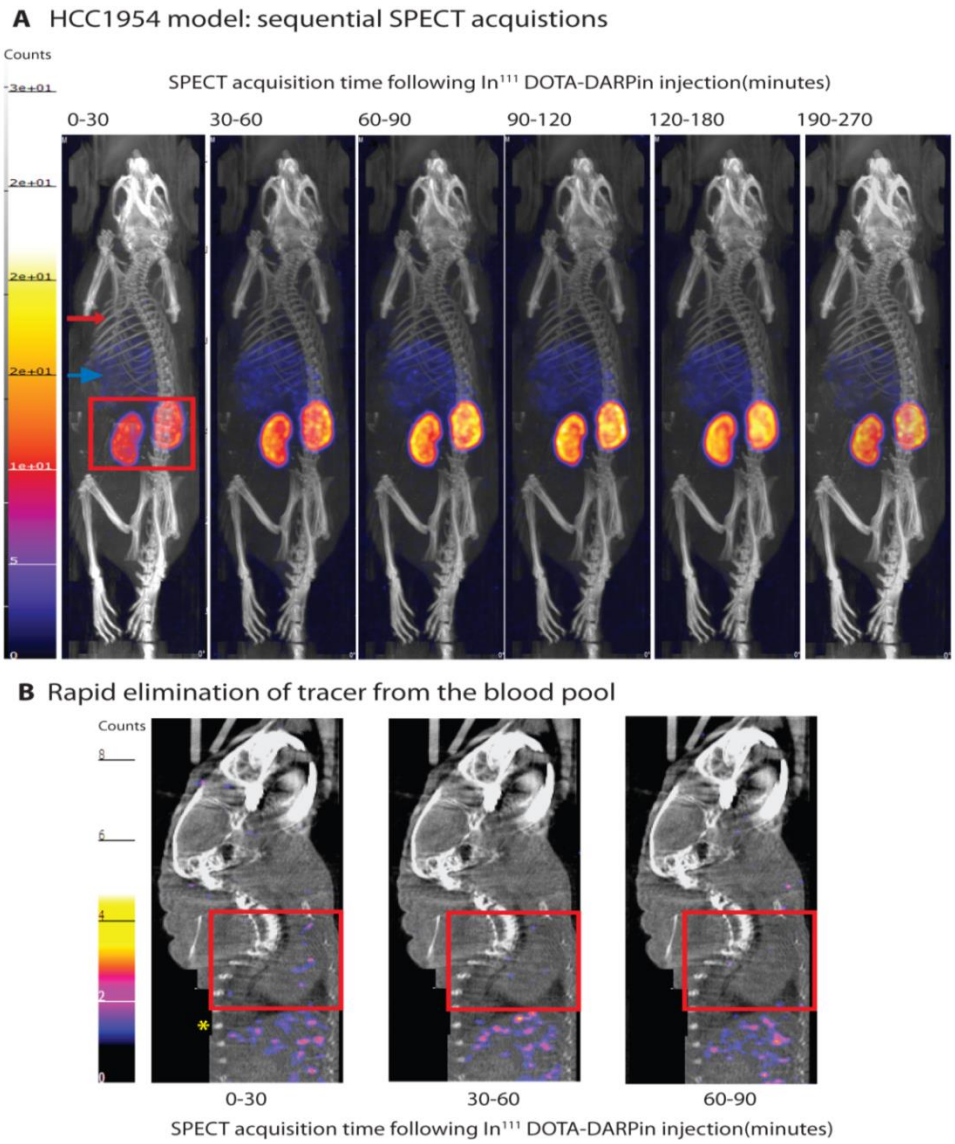


Figure 5-5 Serial SPECT images to establish planned scan time post tracer injection

Sequential SPECT acquisitions were performed in the first scanned HCC1954 mouse (H1, day 33 Luc+ cell line post-injection) in order to establish a suitable time frame for animal imaging following tracer injection.

A. Six acquisitions were performed over 3 hours (60 seconds per frame) following tracer injection. Maximum Intensity projections (images shown) show rapid accumulation and retention of tracer within the kidneys (boxed) and a lesser degree of accumulation within the liver (blue arrow). For the purposes of the figure colour scales have been normalised to facilitate scan comparison, however lung metastases were not visible in the thorax (red arrow) at any time-point on individual image evaluation. Later histological evaluation of fixed tissue confirmed presence of lung metastases in this mouse.

B. The same mouse, sequential sagittal views cropped to highlight cardiac area (red box). Tracer signal is most prominently present within this blood pool on the 0-30 minute scan and

very rapidly clears thereafter. Asterisk indicates location of mouse monitoring pad which has been removed from the images for scan clarity.

5.3.3 Biodistribution of ^{111}*In DOTA-DARPin* tracer in tumour-bearing hosts

Organ biodistribution was calculated using the method described (Chapter 4.5.4). Unfortunately using this method the calculated mouse injected volume was observed to consistently exceed the known total injected volume, for example the calculated volume in harvested organs from mouse H1 was 85.6µl, yet the known injected volume (whole mouse including the tail and gauze used to compress the injection site) was 82µl (calculated from pre- and post-injection syringe weight). The cause of this was unclear but as standards were prepared on both of the scanning days by an experienced radiopharmacist it was thought unlikely to be due to error in standard preparation.

The activity in the whole mouse prior to dissection, residual carcass activity post dissection, and syringe and gauze activity post injection had been measured using the well counter (Capintec, Pennsylvania) on the day of scanning. Therefore these measurements were used to calculate the harvested organ and the carcass activity as a percentage of the total injected activity. The injection volumes delivered to the whole mouse, harvested organs and carcass could then be estimated and the total injection volume delivered to the mouse body derived.

Using this method the calculated % injected dose per gram for all mice in cohort 1 is shown in Table 5.1 below and graphically in Figure 5-6. The low kidney % injected dose per gram seen in mouse H1 can be explained by the relatively higher kidney weight (0.897 grams) in this animal, % injected dose being comparable across the group. Consistent with the visual images no statistically significant difference in lung biodistribution was present between HCC1954 and naïve or HCC1954 and MDA-MB-231 groups ($p=0.52$ and $p=0.83$ respectively, independent samples t test).

Cell line		Lungs	Liver	Spleen	Kidney	Bowel	Heart
HCC1954	H1	1.99	11.34	3.02	55.52	0.89	0.64
	H2	0.59	8.78	2.65	187.90	0.23	0.24
	H3	2.00	10.45	1.66	169.58	0.31	0.69
	H4	8.87	19.13	2.15	247.64	0.48	1.12
	Mean	**3.36**	**12.42**	**2.37**	**165.16**	**0.48**	**0.67**
	SD	**3.73**	**4.60**	**0.59**	**80.33**	**0.29**	**0.36**
	SEM	**1.87**	**2.30**	**0.30**	**40.17**	**0.15**	**0.18**
MDA-MB-231	M1	2.66	10.94	5.38	133.62	0.01	0.23
	M2	2.47	9.05	1.70	124.10	0.98	1.73
	M3	2.20	10.41	3.86	158.92	0.82	1.66
	Mean	**2.44**	**10.13**	**3.65**	**138.88**	**0.60**	**1.21**
	SD	**0.23**	**0.98**	**1.85**	**18.00**	**0.52**	**0.85**
	SEM	**0.13**	**0.56**	**1.07**	**10.39**	**0.30**	**0.49**
Naive	N1	0.28	11.39	2.61	194.41	1.68	1.22

Table 5-1 Cohort 1 ^{111}In DOTA-DARPin biodistribution

Table summarising cohort 1, ^{111}In DOTA-DARPin organ biodistribution. No clear relationship between HER2 metastatic burden and lung biodistribution is apparent in this dataset.

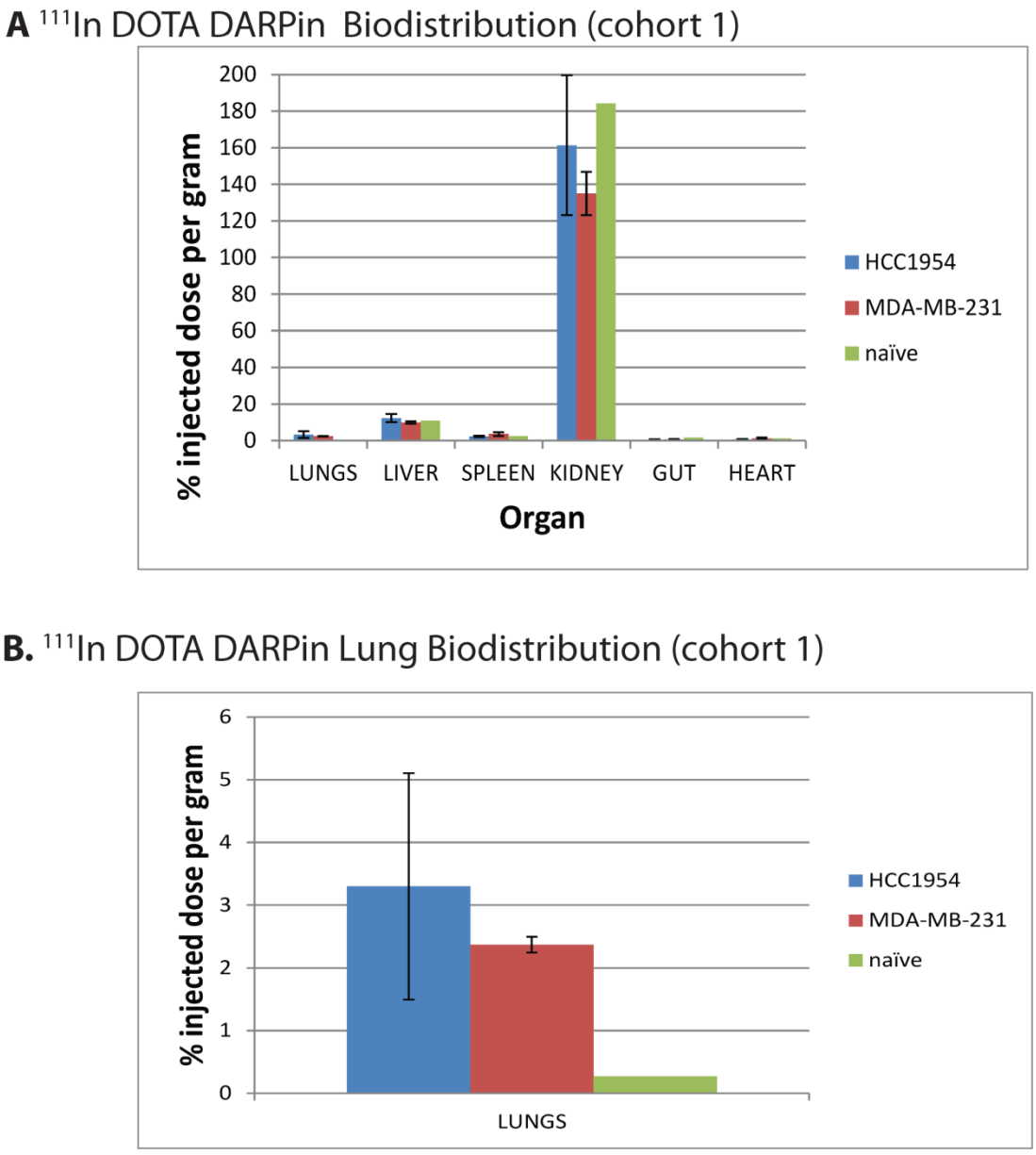


Figure 5-6 Cohort 1 ^{111}In-DOTA-DARPin biodistribution.

Organ biodistribution following ^{111}In-DOTA-DARPin was calculated for all mice in cohort 1 (n=4 in the HCC1954 metastasis group, n=3 in the MDA-MB-231 metastasis group and a single naïve control mouse). In the HCC1954 group, organs were harvested at day 33 (H1), day 30 (H2), day 20 (H3 and H4) after intravenous cell line injection. In the MDA-MB-231 lung metastasis group organ harvest took place at day 33 (n=3) after intravenous cell line injection.

A. ^{111}In-DOTA-DARPin biodistribution for evaluated organs. The mean and SEM are shown for mice in the HCC1954 and MDA-MB-231 metastasis groups.

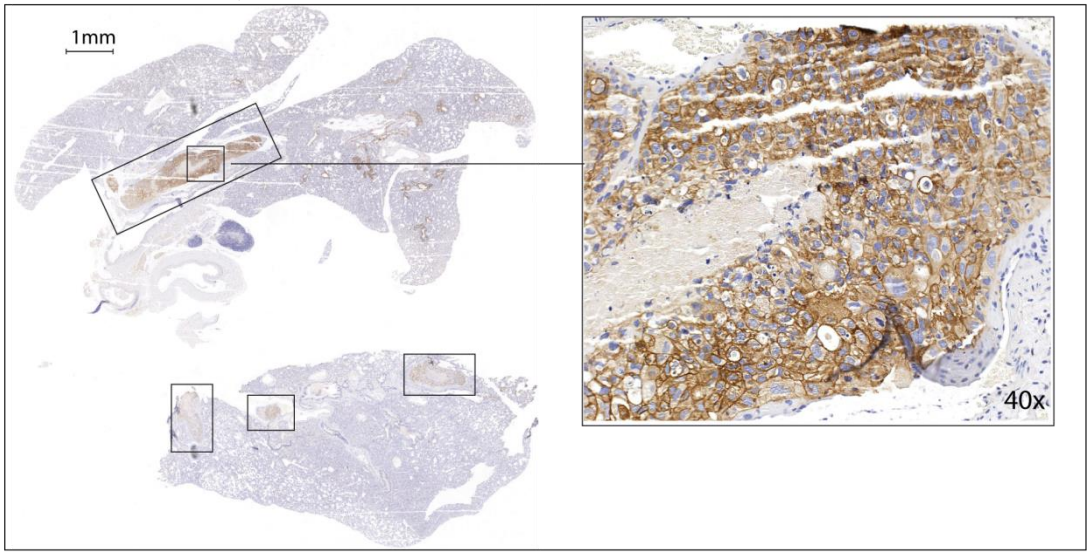
B. Comparison of lung biodistribution across the 3 groups. Mean and SEM shown for mice in the HCC1954 and MDA-MB-231 metastasis groups. No statistically significant difference in lung biodistribution was present between HCC1954 and naïve, or HCC1954 and MDA-MB-231 tumour-bearing hosts (p=0.52 and p=0.83 respectively, independent samples t test).

5.3.4 Histological verification of tumour burden

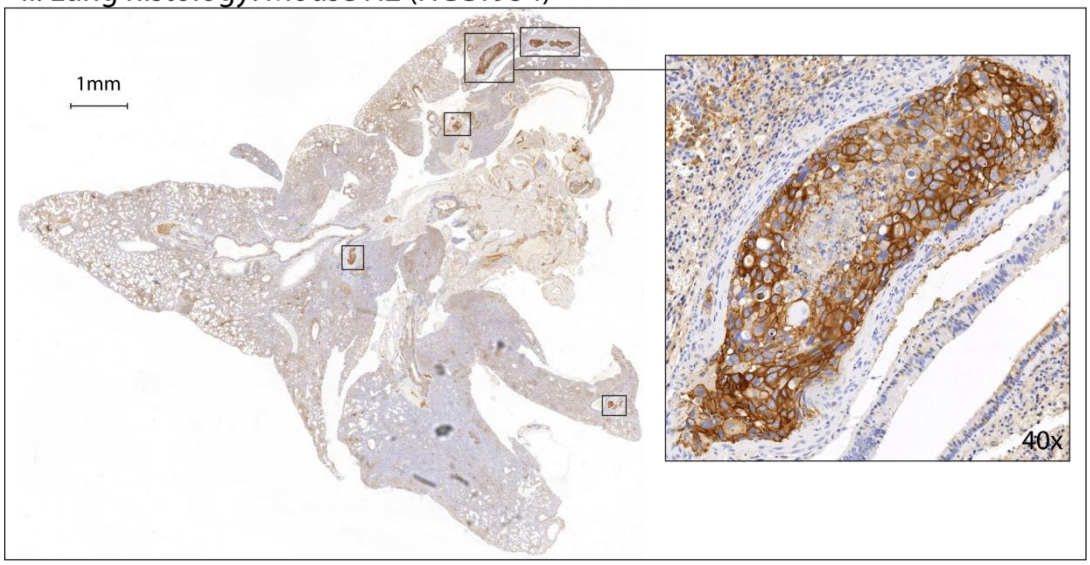
Examination of H&E lung cross-sections confirmed presence of lung metastases in four mice in cohort 1 (n=7). In the HCC1954 tumour-bearing group, multiple clear metastases were demonstrated in mice H1 and H2, and HER2 positivity of these metastases was confirmed using HER2 immunohistochemistry (Figure 5-7 A). Two mice in the MDA-MB-231 tumour-bearing group (n=3) had HER2 negative lung metastases at histology.

Comparison of confirmed metastatic burden with ^{111}In DOTA-DARPin lung biodistribution fails to demonstrate evidence supporting relationship within this first cohort (Figure 5-7 B).

A i. Lung histology: Mouse H1(HCC1954)



ii. Lung histology: Mouse H2 (HCC1954)



B

Model	HCC1954				MDA-MB-231			Naïve
	H1	H2	H3	H4	M1	M2	M3	N1
Lung biodistribution (% injected dose per gram)	1.99	0.59	2	8.87	2.66	2.47	2.2	0.28
Metastasis confirmed	**Yes**	**Yes**	No	No	**Yes**	**Yes**	No	No

Figure 5-7 HER2 immunohistochemistry to confirm presence of metastases in lung section

Lungs from mice in cohort 1 were harvested for histological verification of metastasis at completion of ^{111}In DOTA-DARPin SPECT imaging. In the HCC1954 tumour-bearing group, lungs were harvested at day 33 (H1), day 30 (H2) or day 20 (H3 and H4) after intravenous injection. In the MDA-MB-231 tumour-bearing group, organ harvest took place at day 33 (n=3) after injection. Following radioactivity counting for organ biodistribution, specimens were processed for histological evaluation. Metastases were confirmed in two mice in the HCC1954 group and two mice in the MDA-MB-231 group.

A. Immunohistochemistry confirmed HER2 positive lung metastases (brown circumferential membrane staining) in the HCC1954 group only.

(i) Mouse H1. Low power view (4x, scale 1mm) lung cross-section demonstrating at least 4 Her2+ metastatic tumours, boxed to highlight. A portion of the largest demarcated metastasis measuring 3.33mm, is shown at 40x magnification (scale bar). (ii) Mouse H2. Low power view (4x, scale 1mm) lung cross section demonstrating at least 5 HER2 positive metastases (boxed) in mouse H2. The largest demarcated metastasis measuring 3.33mm, is shown at 40x magnification

B. Table summarising lung biodistribution following ^{111}In DOTA-DARPin and presence of histology confirmed metastases. No relationship between tracer biodistribution and HER2 positive metastatic burden is apparent from this cohort.

5.4 *In vitro* ^{111}In DOTA-DARPin cell binding

To understand the failure of ^{111}In DOTA-DARPin to detect metastases established using the HCC1954 cell line, a tracer-cell pellet binding experiment was performed to evaluate the minimum cell number that can be visualised using the SPECT equipment *in vitro* according to the methods (4.6.1).

Tracer binding to the HCC1954 cell line was confirmed and high correlation was observed between HCC1954 cell pellet size and cell activity (Figure 5.8) (r^2=0.965). In contrast, the HER2-negative MDA-MB-231 cell line could not be visualised even at pellet size of 5 million cells and no activity was present. The minimum number of HCC1954 cells visualised using SPECT imaging was 500,000 cells. Taken together these data indicate that a metastatic tumour below this threshold may potentially explain the failure to visualise the confirmed lung metastases in mice H1 and H2. The cell number contained within each metastasis can be approximated by counting the number of nuclei with in the tumour axial directions and applying the volume formula $((\text{length} \times \text{width})^2)/2$ to give an estimate of total cell number. Using this method the largest demarcated metastasis in mouse H2 (Figure 5.7Aii) might contain as few as 145,800 cells, below the visualisation threshold for a single tumour suggested by the cell pellets.

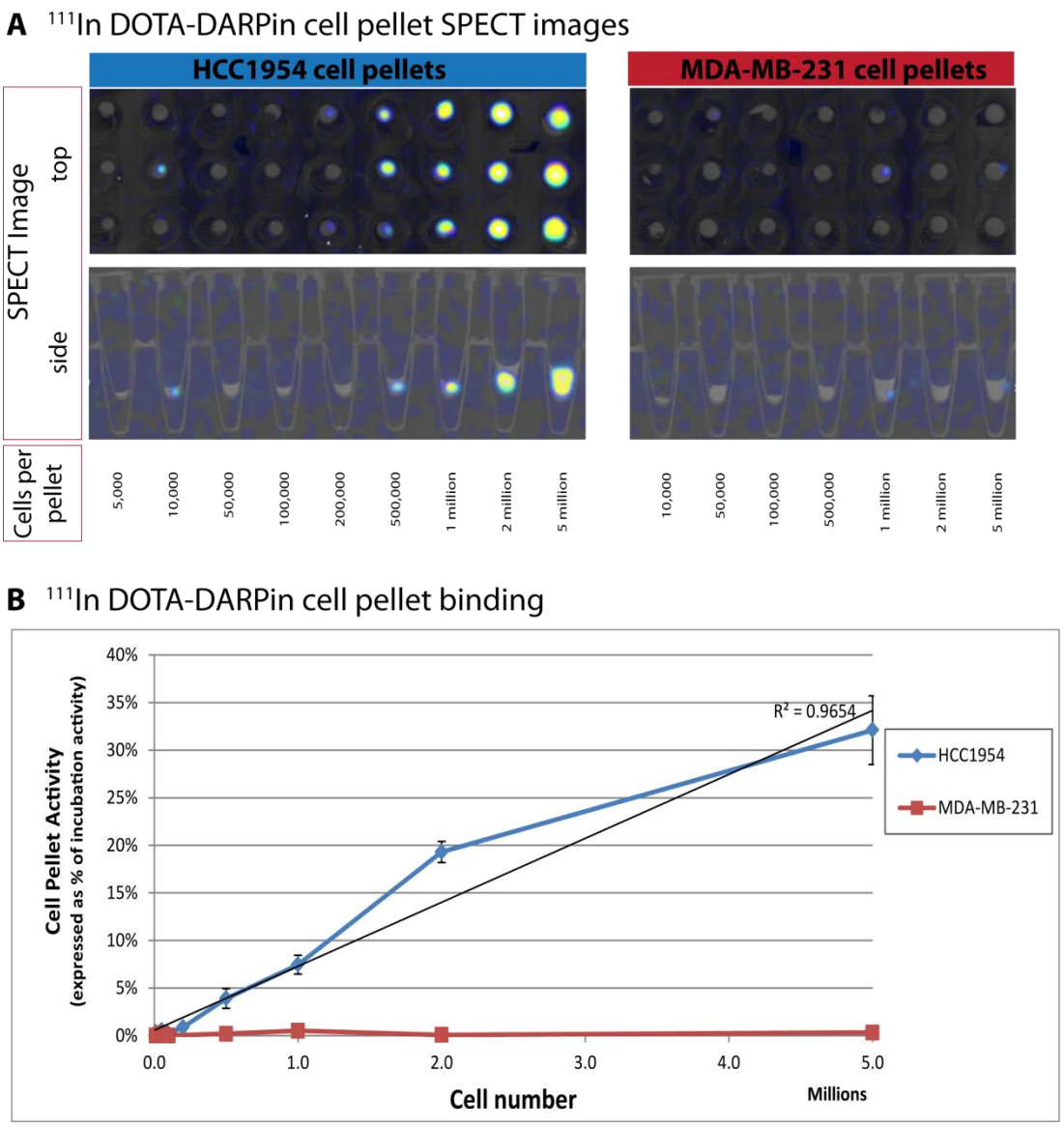


Figure 5-8 ^{111}In DOTA-DARPin cell pellet binding

To determine the minimum cell number detectable by SPECT-CT imaging, serial dilutions of HCC1954 and MDA-MB-231 cells were incubated in triplicate for 30 minutes at 37˚C with ^{111}In DOTA-DARPin tracer. SPECT-CT images were obtained for the washed cell pellets prior to gamma counting to quantify % tracer binding relative to 10 nanomolar incubation standard.

A. SPECT-CT images of MDA-MB-231 and HCC1954 cell pellets following incubation with ^{111}In DOTA-DARPin. Image of the three replicates from the top and side view through central pellet for each pellet size are shown. Only the HER2+ HCC1954 cell line could be visualised at a minimum pellet size of 500,000 cells.

B. Graphical representation of cell pellet activity for the two cell lines. The results represent the mean counts per minute and SEM, expressed as a percentage of the incubation standard, three replicates per pellet. High correlation between HCC1954 cell pellet activity and cell number was observed (r^2=0.97).

5.5 Cohort 2 model preparation

To further evaluate *in vivo* tracer performance in tumours established using the HER+ HCC1954 cell line, a second cohort of 8 mice were prepared for SPECT-CT and biodistribution. Based on the experience with the first cohort, and data from the *in vitro* cell binding experiment, scanning was scheduled at day 46 following cell line injection with the intention of mitigating potential impact of inadequate metastasis formation on tracer performance. A mouse bearing a subcutaneous tumour (D1) and a naïve mouse provided positive and negative controls for the imaging.

Six mice were intravenously injected with 1 million HCC1954 cells on day 0. One mouse experienced respiratory arrest within minutes of injection, however five IV injected mice were suitable for SPECT-CT imaging and biodistribution at the planned time point (mice H5-9). Tumour in positive control mouse D1 was allowed to reach maximum size permitted by the animal licence and measured 12x8mm on day of imaging (day 46).

5.5.1 ^{111}In DOTA-DARPin SPECT-CT imaging and biodistribution

5.5.1.1 ^{111}In DOTA-DARPin Imaging in Cohort 2

SPECT-CT acquisition at 30-60 minutes following ^{111}In-DOTA-DARPin failed to visualise the subcutaneous HCC1954 tumour (Figure 5.10 Ai) despite its large size (Tumour volume 4608mm^3, *ex vivo* weight 500mg on day of imaging). Acquisitions from mice H5, 6 and 7 failed to visualise metastases but showed early renal retention of tracer (images not shown, similar to Figure 5.5).

Biodistribution data from the second cohort confirmed the previously observed renal activity, but no signal was observed with the tumour or lung. Combined biodistribution from cohorts 1 & 2 showed no statistically significant difference in lung biodistribution between groups ($p = 0.525$, one-way ANOVA).

H&E lungs sections failed to confirm presence of metastases in mice H5, H6 & H7, despite the longer interval of 46 days from intravenous cell line injection to day of cull. To exclude the possibility that failure to establish a metastatic burden may be biasing the lung biodistribution, the HCC1954 group from combined cohorts 1 and 2 was split according to histological verification of metastases (yes/no) and differences in lung biodistribution re-evaluated. Once

again no statistically significant difference in lung biodistribution was present between groups separated according to histological verification ($p = 0.516$, one-way ANOVA; (Figure 5.9).

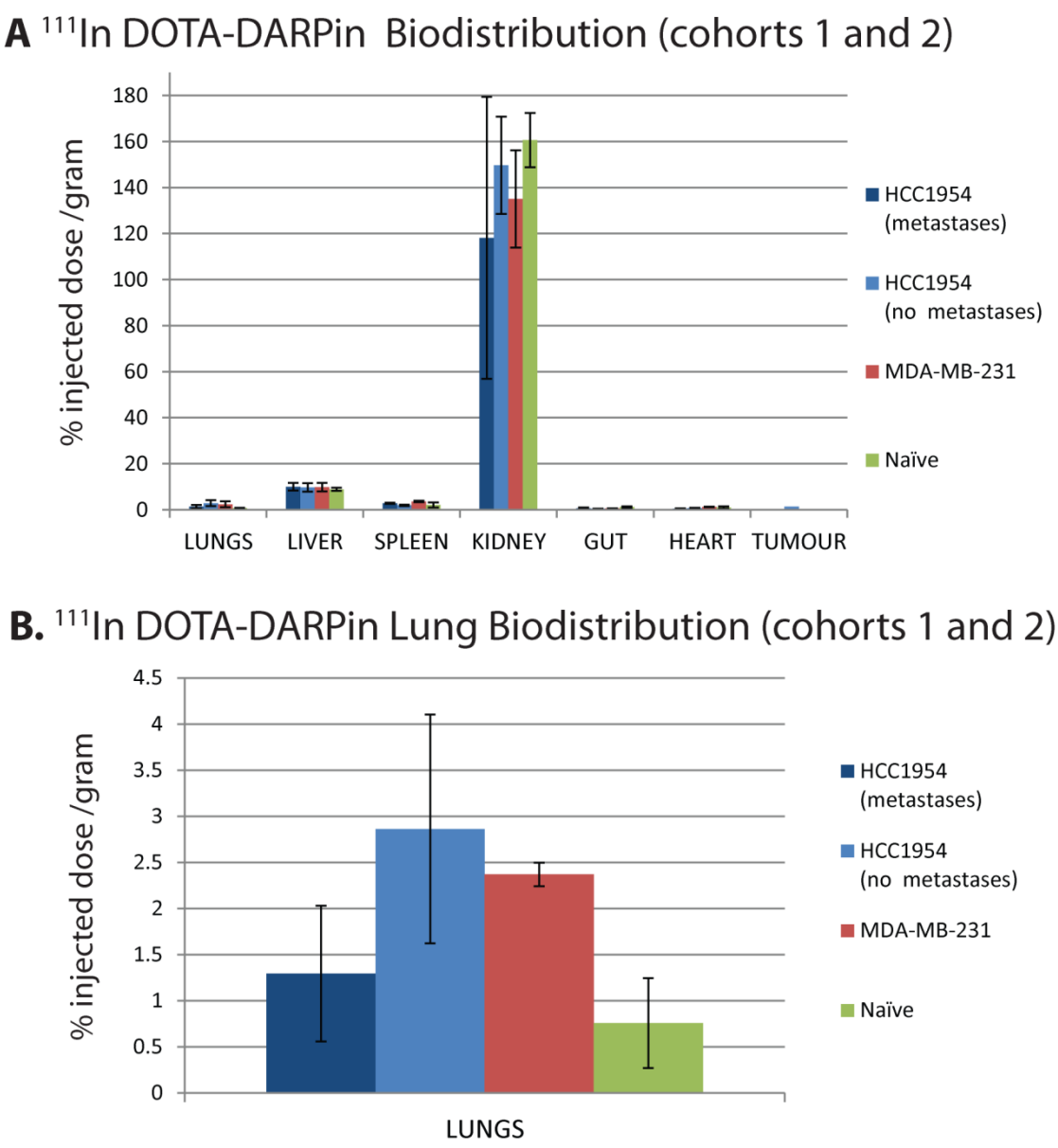


Figure 5-9 ^{111}In-DOTA- DARPin biodistribution from mice in cohorts 1 and 2

A. ^{111}In DOTA-DARPin biodistribution for evaluated organs and excised subcutaneous tumour; combined data from lung metastasis and control mice in cohorts 1 and 2. The HCC1954 injected mice have been separated into those with (H1 and H2) and without (H3-H7) histologically confirmed metastases. Biodistribution of mice in the HCC1954 groups was at day 33 (H1), day30 (H2), day 20 (H3 and H4) and day 46 (H5, 6 and 7) after intravenous cell line injection. Biodistribution of mice in the MDA-MB-231 group is at day 33 (n=3) after intravenous cell line injection. Naïve mice (n=2) provide a non-metastatic control. Organ biodistribution, mean and SEM are shown for the four mouse groups.

B. Comparison of lung biodistribution, mean and SEM shown for mice in HCC1954 with and without metastases, MDA-MB-231 and control groups. No statistically significant difference in lung biodistribution was present between groups ($p = 0.51$, one-way ANOVA).

5.5.2 *In vivo* comparison of ^{111}In-DOTA-DARPin and ^{111}In-CHX-A"-DTPA-trastuzumab tracer performance

Contemporaneous with our experiments, mice bearing subcutaneous HCC1954 tumours had been prepared by a colleague (Dr. Florian Kampermeir, Imaging Sciences) with a view to SPECT-CT evaluation of HER2+ tumours using ^{111}In -CHX-A"-DTPA-trastuzumab. Tumours had been established by injection of 3 million cells injected day 0, and measured 4x4mm (T1) and 5x5mm (T2) on day of tracer injection. SPECT imaging performed at 24 hours (both mice) and 72 hours (T1 only) following tracer injection (Figure 5.10 Ai and ii) clearly demonstrates tumours in both mice in contrast to the images from the ^{111}In-DOTA-DARPin imaged mouse (D1). The *ex vivo* ^{111}In -CHX-A"-DTPA-trastuzumab biodistribution supports the tumour visualisation seen on the SPECT images (Figure 5.11). T2 was culled at 24 hours following tracer injection and relatively higher splenic activity is consistent with expected antibody biodistribution at this early time point following tracer injection (116, 117).

Excised tumours from mice D1, T1 and T2 weighed 500.0mg, 108.6mg and 60.0mg respectively. As expected H&E sections demonstrated central necrosis was more marked within the larger D1 tumour, however HER2 positivity was confirmed in all three tumours (Figure 5.10B).

Given the superior imaging performance of ^{111}In-CHX-A"-DTPA-trastuzumab observed with mice bearing subcutaneous HCC1954 tumours, the two remaining HCC1954 tail vein injected mice from cohort 2 (H8 and H9) were imaged with this tracer. SPECT acquisitions performed at 24 and 72 hours following tracer injection (day 50 following cell line injection) failed to visualise lung metastases in either mouse (images not shown) and biodistribution similarly provides no lung activity signal compared to mice T1 and T2 with subcutaneous tumours (Figure 5.11). However despite the 50 day interval from HCC1954 injection, no metastases were identified by subsequent lung histology SPECT images and therefore failure to visualise metastasis cannot be attributed to *in vivo* tracer performance. It is unclear why lung metastases could not be established in this tail vein injected cohort but colocation of BLI imaging to track tumour burden prior to SPECT scanning and organ harvest would be desirable prior to any further evaluation of the ^{111}In-CHX-A"-DTPA-trastuzumab tracer in metastatic xenograft models.

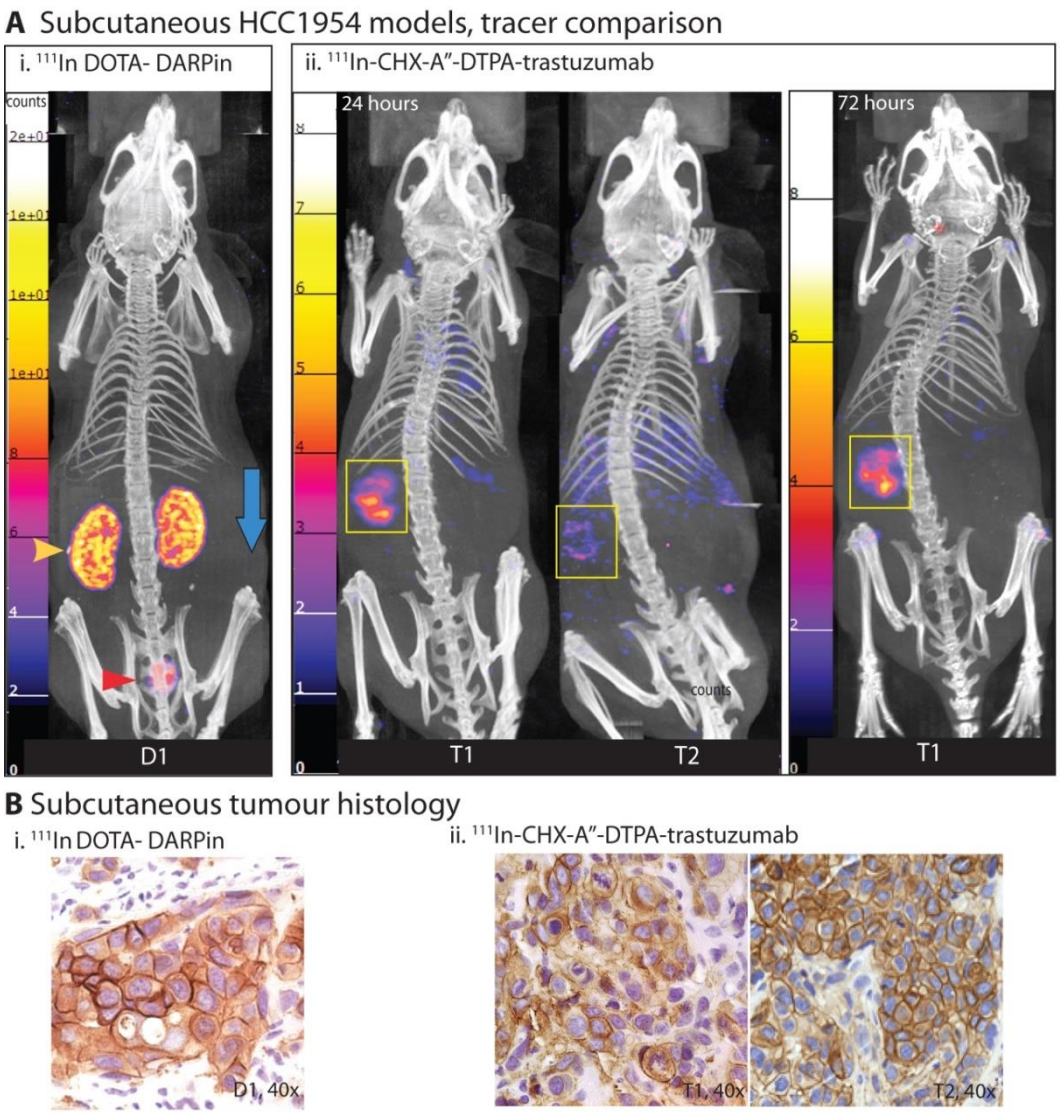


Figure 5-10 Subcutaneous model ^{111}In DOTA-DARPin and ^{111}In -CHX-A"-DTPA-trastuzumab tracer comparison

In vivo imaging performance of ^{111}In DOTA-DARPin and ^{111}In-CHX-A"-DTPA-trastuzumab tracer performance was compared in mice bearing HCC1954 subcutaneous tumours. Mouse D1 underwent SPECT-CT imaging 30 minutes following injection ^{111}In DOTA-DARPin (46 days after injection of 1 million cells, excised tumour weight 500mg). ^{111}In-CHX-A"-DTPA-trastuzumab imaged mice, T1 and T2, underwent SPECT-CT imaging at 24 hours following tracer administration, 23 days after injection of 3 million cells. Imaging was repeated in mouse T1 at 72 hours. Excised tumours weighed 108.6mg and 60mg respectively for mouse T1 and T2.

A. SPECT-CT Maximum intensity projection views. (i) ^{111}In DOTA-DARPin SPECT image acquisition. Renal excretion of tracer is clearly visualised within kidney (yellow arrow) and bladder (red arrow) but the large palpable tumour (blue arrow) cannot be discerned. (ii) ^{111}In -CHX-A"-DTPA-trastuzumab SPECT images clearly visualise the palpable subcutaneous tumours (yellow boxes).

B. HER2 immunohistochemistry confirming presence of HER2+ tumour cells in both the ^{111}In DOTA-DARPin (i) and ^{111}In -CHX-A"-DTPA-trastuzumab (ii) imaged mice.

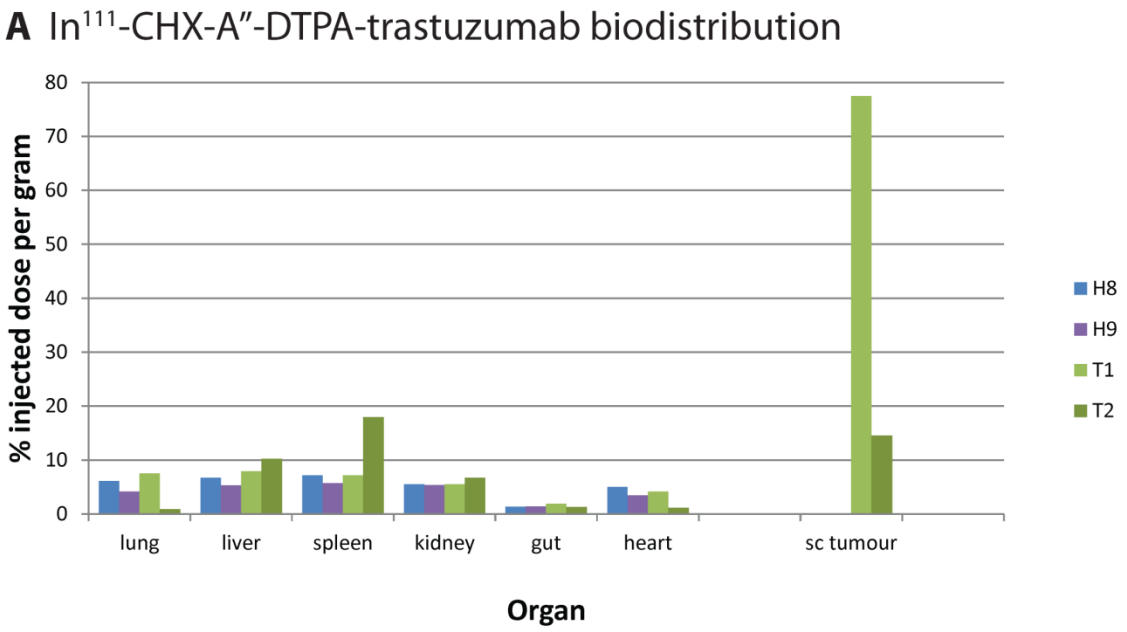


Figure 5-11 ^{111}In -CHX-A"-DTPA-trastuzumab organ biodistribution

Biodistribution for ^{111}In-CHX-A"-DTPA-trastuzumab imaged mice. Mice H8 and H9 underwent SPECT imaging at day 50 following intravenous HCC1954 cell line injection; T1 and T2 underwent SPECT after establishment of HCC1954 subcutaneous tumours (volumes 128 and 312mm^3 respectively). Organ harvest took place immediately following final SPECT image acquisition at 72 hours (H8, H9, T1) and 24 hours (T2).

Lung biodistribution does not discernibly differ between the subcutaneous and tail vein models (organ harvest at 72 hours mice H8, H9, T1) and is consistent with the absence of histologically verified lung metastasis in any animal. Relatively high biodistribution in histologically verified subcutaneous tumours is consistent with the SPECT image acquisitions which demonstrate clearer tumour visualisation in mouse T1 than in T2 (figure 4.5).

5.6 *In vitro* ^{111}In-CHX-A”-DTPA-trastuzumab cell binding

In view of the superior *in vivo* performance of ^{111}In-CHX-A”-DTPA-trastuzumab for subcutaneous tumour imaging, the cell pellet imaging was repeated using cells incubated with this tracer to evaluate the minimum cell number that can be visualised using SPECT *in vitro*. ^{111}In-CHX-A”-DTPA-trastuzumab binding to the HCC1954 cell line was confirmed and activity correlated highly with cell number (r^2=0.996). The minimum number of cells that could be visualised was 100,000, in contrast to 500,000 using ^{111}In DOTA-DARPin. MDA-MB-231 cells could not be visualised even at pellet size of 5 million cells and no activity was present (Figure 5.12).

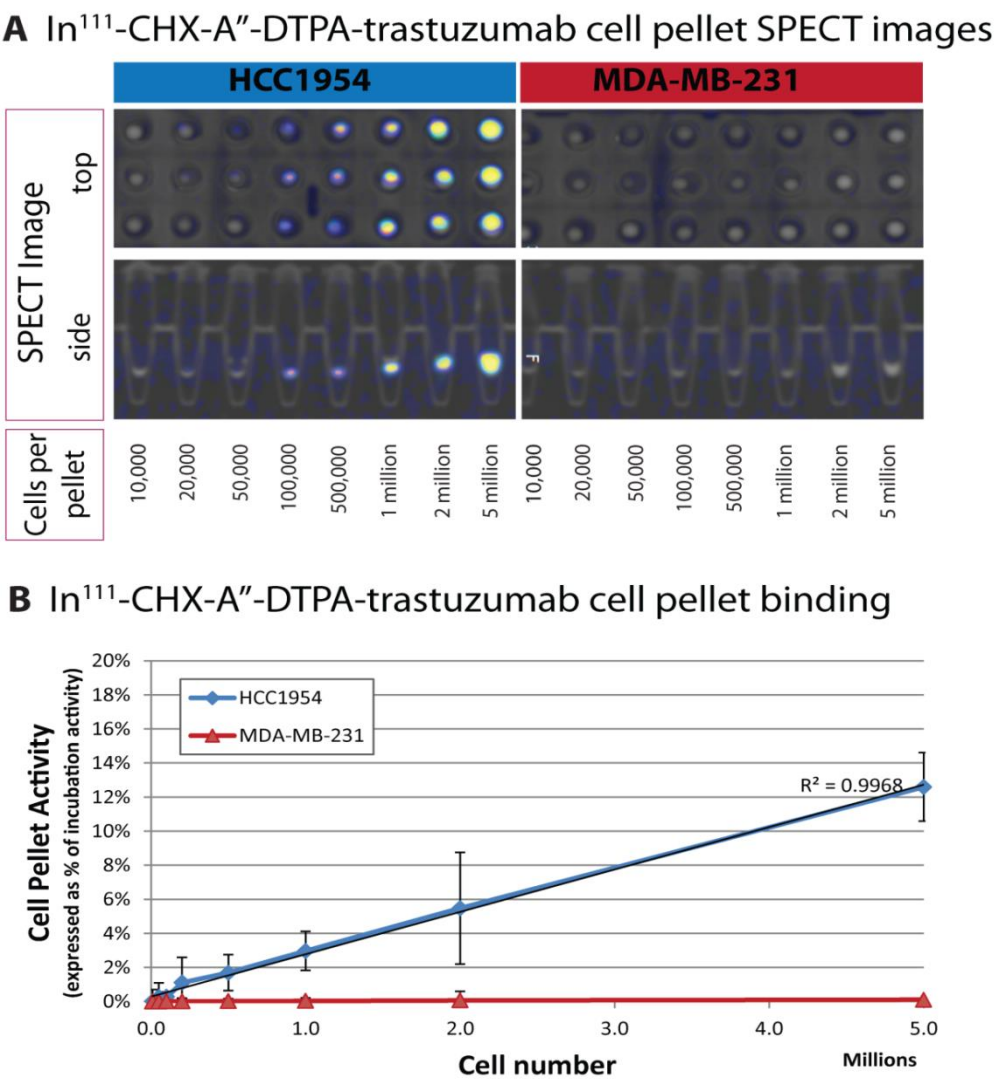


Figure 5-12 ^{111}In-CHX-A"-DTPA-trastuzumab cell pellet binding

To determine the minimal HER2+ cell number detectable by SPECT imaging serial dilutions of HCC1954 cells were incubated in triplicate with ^{111}In-CHX-A"-DTPA-trastuzumab for 30 minutes at 37˚C. SPECT-CT images were obtained for washed cell pellets prior to gamma counting to quantify % tracer binding relative to the 10 nanomolar incubation standard. MDA-MB-231 was used as the HER2-negative control.

A. SPECT-CT images of cell pellets following incubation with ^{111}In-CHX-A"-DTPA-trastuzumab. Maximum Intensity Projection (top) views of the three replicates and sagittal view (side) through central pellet for each pellet size are shown. HER2+ HCC1954 cell pellets are clearly visualised, minimum cell number 1,000,000 cells. No signal is present in MDA-MB-231 cell pellets.

B. Graphical representation of cell pellet activity. The results represent the mean counts per minute and SEM, expressed as a percentage of the incubation standard, three replicates per pellet. High correlation between HCC1954 cell pellet activity and cell number is observed (r^2 =0.997).

5.7 *In vivo* pellet experiment

In view of discrepancy between the *in vitro* cell pellets images and the *in vivo* results, an experiment to confirm that cells incubated with tracer *in vitro* could be then visualised *in vivo* was performed.

A single mouse was injected with cell suspensions that had been pre-incubated with either ^{111}In-CHX-A"-DTPA-trastuzumab or ^{111}In DOTA-DARPin. For each tracer 5 million cells were incubated with tracer at 10nM concentration for 30 minutes at 37˚C, as previously described (4.6.1). Following incubation cells were washed twice with sterile PBS and re-suspended in 250µl sterile PBS for injection. Measured activity was 100KBq for both preparations.

2 million cells (100 µl) of ^{111}In-CHX-A"-DTPA-trastuzumab and ^{111}In DOTA-DARPin cell preparations, and 0.5 million cells (25µl) of the ^{111}In-CHX-A"-DTPA-trastuzumab cell preparation were injected at separate subcutaneous locations in a single mouse prior to SPECT-CT image acquisition. All pellets could be visualised (Figure 5.13). Taken together this suggests that failure of delivery of tracer to tumours may have contributed to the poor *in vivo* performance of ^{111}In G3 DOTA-DARPin.

To ensure that tracer could be delivered to tumours *in vivo*, the vascularity of the excised subcutaneous tumours from the second mouse cohort was quantified. These tumours were previously imaged using ^{111}In-CHX-A"-DTPA-trastuzumab (mouse T1 and T2) and ^{111}In G3 DOTA-DARPin (mouse D1). Despite the disparity in tumour size (ex vivo weight T1 = 108mg, T2 =60mg D1=500mg) no statistically significant difference in the vessel count was present between tumours (Figure 14). Taken together it is more likely that the rapid clearance of ^{111}In G3 DOTA-DARPin tracer from the circulation, rather than differences in intra-tumoral tracer delivery, prevented adequate tumour uptake.

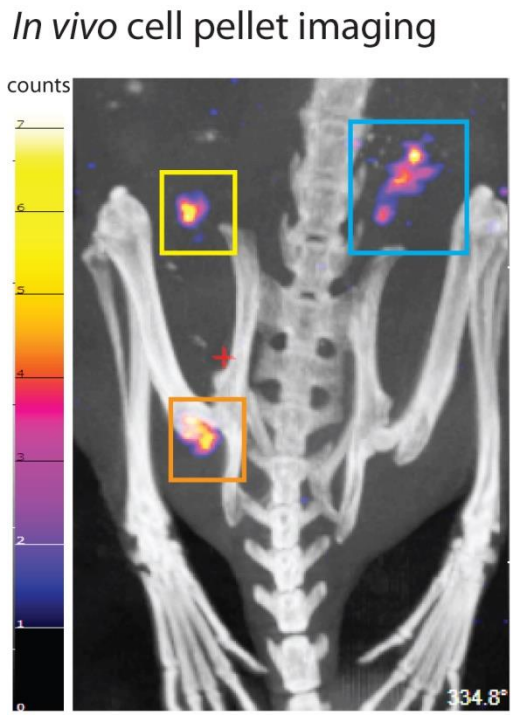


Figure 5-13 *in vivo* cell pellet imaging

SPECT acquisition from a single mouse injected with HCC 1954 cell pellets incubated for 30 minutes with ^{111}In DOTA-DARPin (2 million cells, blue box) or ^{111}In-CHX-A"-DTPA-trastuzumab (2 million yellow box, 0.5 million orange box). Cell pellets are clearly present on SPECT acquisition using both tracers. This suggests that failure to image the tumour using the ^{111}In DOTA-DARPin at SPECT imaging may be due to inadequate tracer delivery to tumour in vivo.

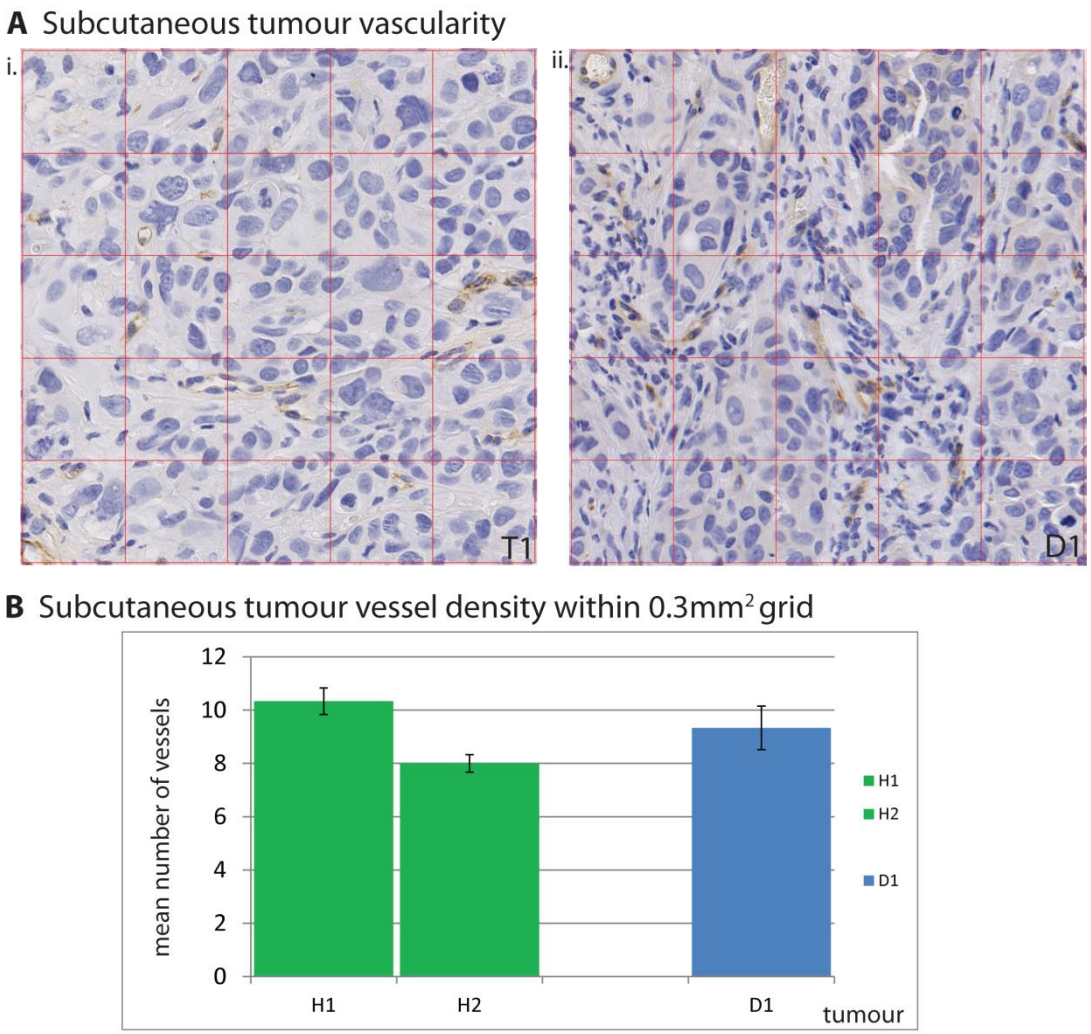


Figure 5-14 Subcutaneous tumour vascularity

A. CD31 immunohistochemistry was performed on 3µm sections from HCC1954 subcutaneous tumours. Vascular endothelium stains brown. (i) Subcutaneous tumour section 0.3mm² grid resected following SPECT imaging with *^{111}In-CHX-A"-DTPA-trastuzumab (mouse T1). (ii) Subcutaneous tumour section 0.3mm² grids resected following SPECT imaging with ^{111}In DOTA-DARPin (mouse D1).*

B. The most intense vascular areas (hotspots) were selected subjectively from each tumour section and vessel count (Mean and SEM) quantified by counting the number of vessels present within three 0.3mm² grids placed within these areas. There was no statistically significant difference in the vessel count between mice (p = 0.345, one-way ANOVA).

5.8 Conclusions

The data presented in this chapter demonstrates that ^{111}In-DOTA-DARPin binds to HER2+ HCC1954 cells *in vitro* and SPECT imaging permits discrimination between HCC1954 and HER2-negative MDA-MB-231 cell pellets. However, in contrast to radiolabelled trastuzumab, this tracer failed to discriminate between HER2+ and HER2-negative tumour models *in vivo*. The (18)F-Z(HER2)(:342)-Affibody has been reported to successfully targeted HER2+ lesions in the lung permitting detection of metastases as early as 9 weeks following cell line injection (132). It is possible that the small size of pulmonary metastases in our metastatic model accounts for failure of the DARPin tracer to visualise lung lesions. However small target lesion size does not explain the failure to visualise the large subcutaneous control tumour with confirmed HER2 expression and vascularisation *ex vivo*, particularly given the tracer binding that was demonstrated to the HCC1954 cell line *in vitro*. In the context of previously reported high tumour-to-blood and tumour-to-normal tissue ratios permitting same day ^{111}In-DOTA-DARPin SPECT imaging of subcutaneous xenografts (BT474 Her2+ breast cell line) (141) the poor *in vivo* performance of this tracer was unexpected. The SPECT images indicated very rapid tracer clearance from the blood pool and accumulation of the radioactivity in the kidney on the early sequential imaging acquisitions. This is a well-recognised challenge for proteins with a molecular weight below 60 kDa (126, 143). Specifically a short half-life of less than 3 minutes has previously been reported with the G3 DARPin (4). Therefore rapid clearance from the circulation, precluding tracer binding to target tumour sites is a plausible explanation for the discrepancy between the observed *in vitro* and *in vivo* performance.

In the current experiments ^{111}In DOTA-DARPin protein was from the same preparation used in the experiments performed at UCL (provided courtesy of Dr R Goldstein), radiolabelling was performed by the same radiopharmacist and the same ^{111}In DOTA-DARPin tracer dose was injected for scanning. Therefore differences in tracer preparation are unlikely to account for the results observed. SPECT imaging was able to discriminate HCC1954 cell pellets incubated with ^{111}In DOTA-DARPin and therefore intrinsic differences between the BT474 and HCC1954 cell lines are also unlikely to explain the tracer failure *in vivo*. The UCL group used female BALB/c mice (Charles River) or SCID-beige mice (Charles River) and as is possible that factors related to biology of difference mouse strains impact on tracer performance. However our bio-distribution data failed to provide a signal of *in vivo* targeting ability and the high renal bio-distribution is comparable to that reported by the UCL group (232.0±24.1 %ID/g), mirroring the very high renal uptake demonstrated on their SPECT image acquisitions (141).

In view of the previously reported efficacy of this tracer in discriminating HER2+ tumours established from BT474 breast cancer cell line, a direct comparison between HCC1954 and

BT474 subcutaneous xenograft models and a HER-negative control would be recommended as a minimum before progressing clinical applications of this tracer. The reason for failure of the HCC1954 cell line to establish metastasis in the second tail vein injected cohort is unclear and it is regrettable that it was not possible for us to use BLI to indicate presence of lung metastases prior to SPECT. Although BLI reflects the number of metabolically active tumour cells rather than a volumetric measurement of tumour mass, longitudinal BLI offers the ability to efficiently confirm establishment of metastases without animal sacrifice. Co-location of BLI and SPECT imaging facilities within the same BSU would facilitate selection of animals with a confirmed metastatic burden permitting more efficient evaluation of tracer performance and would be desirable prior to any future work using this or other HER2 targeted tracers in metastatic xenograft models .

6 TNPET01 Study set up and methodology

6.1 Introduction

To address the unresolved questions regarding tracer selection, scan acquisition and interpretation for response assessment I designed an imaging feasibility study for evaluation of early PET response in the TNBC phenotype. To confirm this as an area of unmet need, in preparation of the study protocol and funding application, I performed a detailed review of over 200 trials listed on clinicaltrials.gov database under the search terms PET, FDG or FLT and breast and identified no trials targeted at the TNBC subset. 27 of the 31 listed response evaluation trials used only a single tracer in unselected breast cancer (17 trials) or receptor positive breast cancer subset (10 trials). The proposed treatment was largely undefined and timing of response was variable. Only 4 listed studies evaluated both FLT and FDG tracers, all in unselected heterogeneous breast cancer.

This chapter reports the development of the full study protocol and the methodologies used for the imaging research analyses. The design of the PET imaging protocol recognises the uncertainties surrounding optimal image acquisition, SUV reporting, and the requirement for dynamic evaluation to optimise PET response assessment for future prospective trial design in this phenotype. The trial was designed with two parts to include a baseline test-retest assessment for each tracer to measure repeatability and assess the confidence with which any change in tracer uptake observed during treatment can be interpreted as related to therapy within the TNBC population. Included within this chapter is the set-up work completed in parallel with protocol writing, including securing funding, regulatory approvals and electronic case report form (eCRF) development.

6.2 Protocol Aims

The purpose of the trial was to assess the repeatability and utility of PET-CT functional imaging using FLT and FDG tracers to predict TNBC tumour response to standard neoadjuvant chemotherapy.

6.2.1 Primary Objectives

- Part A: To measure PET scan repeatability using FDG and FLT tracers and to determine the optimal tracer for response evaluation in part B.

- Part B: To evaluate early PET imaging using either FDG or FLT tracer as methods for evaluating systemic therapy response in primary triple negative breast cancer with respect to MRI response at 3 cycles.

6.2.2 Secondary Objectives

- To relate changes in tracer uptake to final clinical and pathological (RCB) response
- To relate changes in tracer uptake to biopsy derived biomarkers
- To obtain exploratory data relating to tracer kinetics (i.e. Ki, MRGlucose)
- To obtain exploratory performance estimates for early PET and early MRI scans at 1 cycle to predict subsequent clinical and pathological (RCB) response

6.3 Trial Design

The study was designed as a single centre randomised phase II trial of FLT and FDG tracers to predict systemic therapy response in ER- and HER2- breast cancer compared to standard MRI imaging response and biopsy derived biomarkers. The study was designed with two parts according to the overview summarised in Figure 6.1. Accrual to Part A has completed and Part B recruitment continues.

In Part A (10 patients) participants underwent two baseline PET-CT scans to measure the repeatability of PET scan SUV measurements using FDG and FLT tracers. Participants underwent one PET-CT scan with either FDG or FLT to compare early PET response (day 17±3 post cycle 1) with standard MRI response at 3 cycles. The protocol mandated criteria for Part A to B tracer transition were static scan repeatability of within ±15% and evidence of SUV drop of >20% in at least 50% MRI defined responder for either tracer to progress to part A. In the event of equivalent results within the limits set for repeatability, the Trial Steering Committee (TSG) were required to decide which tracer performed best overall and should be taken forward to part B. In the event that neither tracer met the repeatability requirement for at least one SUV measure the study would terminate without progression to Part B.

In part B a total of 15 further patients were required to undergo PET-CT imaging, once prior to commencing chemotherapy and again at day 14 to 21 post cycle 1 using the single tracer selected by the TSG according to end of Part A criteria. Recruitment to part B continues.

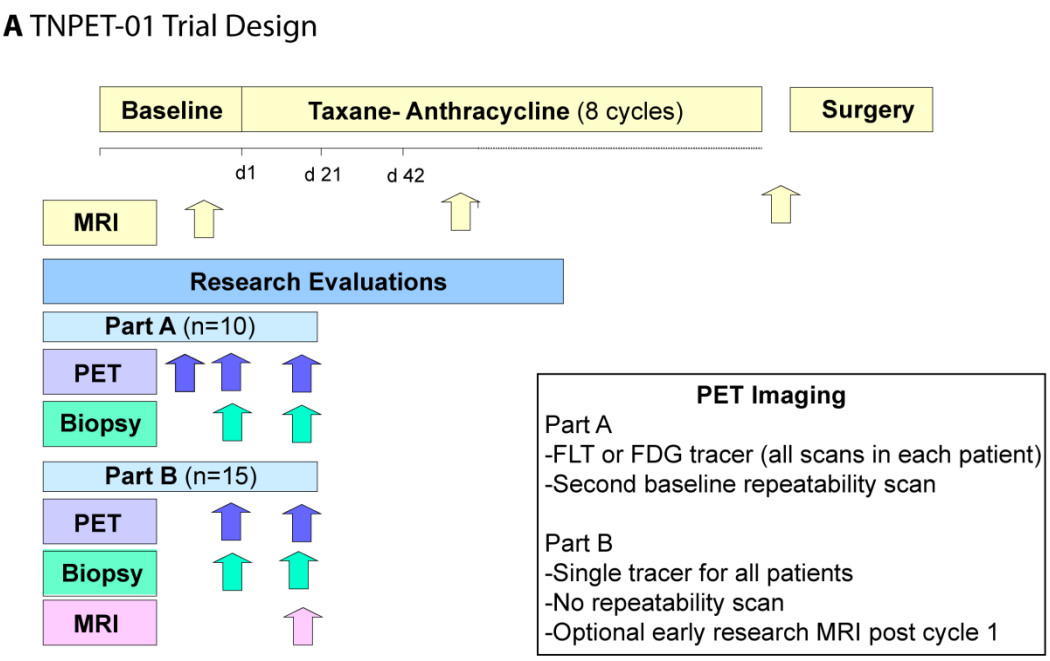


Figure 6-1 TNPET-01 study design

Schematic diagram summarising the TNPET-01 design. For either tracer to be considered for Part A to B transition they had to meet protocol mandated SUV repeatability and post-cycle 1 SUV change criteria. A *single tracer was selected for progression to Part B by the Trial Steering Committee following review of the Part A data.*

In addition to the PET imaging evaluations, study participation (parts A and B) required patients to undergo a research core biopsy performed at three time points (prior to chemotherapy, following their day 17±3 PET scan and through the definitive resection specimen at the time of surgery) and permitted an optional research blood sample. In part B only, an optional early MRI evaluation was included to take place contemporaneously with the second PET scan. Due to the sequential use of taxanes and anthracycline chemotherapy in standard neoadjuvant chemotherapy, eventual pathological response will be influenced by the later chemotherapy cycles of a different class to that reported by an early PET-CT. Consequently the selected primary endpoint of the full study is feasibility of early PET-CT to determine standard MRI RECIST response after 3 cycles and pathological endpoint will be reported as a secondary endpoint.

6.4 Study population

Eligible female patients age 18-70 with stage II-III biopsy proven primary ER- (Allred <3) and HER2- (IHC 0 or 1+, or IHC 2+ and FISH non-amplified (ratio of Her2 to chromosome 17 of more

than 2.0) breast tumours for whom neoadjuvant chemotherapy had been recommended were recruited to the study. Pregnant or breast feeding patents and those with diabetes or serious medical conditions likely to compromise ability to complete study imaging and translational research biopsies were excluded.

All patients were given a patient information sheet (PIS) approved by the Westminster Research and Ethics group (REC reference 11/L0/1492). Written informed consent was obtained prior to any study specific procedures.

6.5 PET scanning protocol

6.5.1 PET-CT scan acquisition

Participants consenting to Part A only were randomised in a 1:1 ratio to receive FDG or FLT tracer to balance tumour and patient factors to avoid bias in tracer performance assessment. Tracer allocation was stratified for tumour nodal involvement (yes/no) and performed centrally via the Institute of Cancer Research Clinical Trials Statistical Unit (ICR-CTSU). All patients consenting to Part B were imaged using the progressing FDG or FLT tracer. In all patients PET-CT scanning was performed prior to commencing chemotherapy and at day 17±3 days following the first chemotherapy cycle. In part A only participants underwent a second baseline PET-CT scan to assess repeatability a minimum of 24 hours after the first.

PET imaging was performed using a GE Discovery VCT 64 slice PET-CT scanner prior to the PET centre refurbishment (September 2013) or using a GE Discovery 710 64 slice PET-CT scanner subsequently. The injected dose was calculated by measurement of the 18F-FDG activity in the syringe before and after the injection. Patient height, weight, fasting time, serum glucose prior tracer injection and blood pressure prior to and following the final acquisition were recorded. For each scan participants underwent a maximum of 3 CT components for attenuation correction and localisation of the dynamic and static scan acquisitions (2 FOV, 140KVp, mAs=28-42, coll=40mm, pitch=1.375). Static PET scans were reconstructed with Iterative VUE Point (OSEM (ordered subset expectation maximisation)-based algorithm; 20 subsets, 2 iterations, 6.0mm Gaussian filter, matrix 128 x 128). PET scans on the Discovery 710 scanner were reconstructed with VUE Point FX (time-of-flight incorporated into VUE Point HD); (24 subsets, 2 iterations, 6.4mm Gaussian filter, matrix 256 x 256).

6.5.1.1 FDG PET

Patients received an intravenous injection of maximum 200MBq FDG (MetaTracer FDG solution for injection, Siemens PLC, UK) whilst on the scanner couch. A dynamic acquisition (1x10, 10x5, 6x10, 3x20, 78x60secs) was performed over the breast primary. The rate of ^{18}F-FDG is taken uptake and retention varies according to tissue but in malignant tumors the accumulation of ^{18}F seldom reaches a plateau by 2 h after injection (144, 145). Therefore to address unanswered questions concerning optimal interval following FDG tracer injection for image acquisition the dynamic imaging was followed by three static scans over the breast and axilla area at 90, 120 and 180 (10 minutes per bed position). Patients were able to get up between scan acquisitions and urinate as required. Images were acquired in 3D mode, 1 bed position for the dynamic and 1 or 2 bed positions of 10 minute duration for the static views.

Patients who gave additional consent had a single venous sample taken during the imaging studies at 60 minutes to record the radioactivity within the blood following injection of the radioactive FDG IMP.

6.5.1.2 FLT PET

FLT was synthesised by the PET centre radio pharmacy (King's College London/Guys and St Thomas' NHS Foundation Trust, London). Patients received an intravenous injection of maximum 200MBq FLT. A dynamic acquisition (1x10, 10x5, 6x10, 3x20, 78x60secs) was performed over the breast primary followed by a single static scan over the breast and axilla area at 90 minutes. Images were acquired in 3D mode, 1 bed position for the dynamic and 1 or 2 bed positions of 10 minute duration for the static views.

Patients who gave consent had 5 venous samples taken during the imaging studies at 5, 10, 30, 60 and 100 minutes to record the radioactivity within the blood following injection of the radioactive FLT IMP and to permit the analysis of metabolites (the glucuronide form of FLT) for dynamic scan data correction. These samples were taken via a second venous cannula inserted into the opposite arm to the arm into which tracer was administered and linked to a sterile cannula tubing. Samples were analysed using a solid phase extraction chromatography technique to separate FLT from the FLT glucuronide for dynamic scan interpretation with metabolite correction according to previously described methods (146).

Twenty-four hours after completion of each FLT scan patients were telephoned to document any toxicities experienced.

6.5.2 Radiation effective dose

A research ARSAC certificate was obtained for this study. Table 4.2 summarises the radiation effective dose for each participant receiving a maximum of three PET-CT scans (Part A) or two PET-CT scans (Part B). It was anticipated that the majority of study patients would receive therapeutic adjuvant radiotherapy dose of at least 40-50Gy to the breast or chest wall ± 16Gy tumour bed boost that far exceeds the research PET imaging dose.

	FDG	FLT
PET component (maximum 200MBq)	4mSv	6.5mSv
CT component	4.5mSv *	2mSv
Total effective dose per session	8.5mSv	8.5mSv
Part A total effective dose (3 sessions)	25.5mSv	25.5mSv
Part B total effective dose (2 sessions)	17mSv	17mSv

Table 6-1 Radiation effective dose for study participants

*Radiation effective dose summarised for PET and CT imaging components. *CT dose in the FDG groups reflects additional CT acquisitions required in the event participants leave the scanning couch between the three static acquisitions at 90, 120 and 180 minutes.*

6.5.3 PET-CT scan interpretation methods

All scan acquisitions were imported into HERMES workstation (Hermes Medical Solutions, Stockholm, Sweden) for research SUV analysis. Suitability of the research scans for Part A and B analysis was confirmed through visual inspection of the static scan by an experienced nuclear medicine physician (Dr Michael O'Doherty, MD or Dr Sally Barrington, SB). Presence of abnormal uptake (where applicable) was defined as uptake not related to expected physiological uptake and higher than adjacent normal tissue. In the event that areas of previously unknown tumour involvement were identified on the pre-treatment research PET scan, a clinical report was issued and the findings were fed back to the clinical team. Scans were subsequently anonymised for research analysis.

Static scan evaluation was performed using the HERMES HybridViewer Software version 1.4C. All VOIs were defined by a single operator, Jennifer Glendenning (JG). At each scan time point, tumour volumes of interest (VOIs) were generated by manually defining a constraint in the axial, sagittal and coronal planes around each area of interest. A tumour isocontour was then determined within this constraint using a 40% threshold of the maximum uptake to reduce operator bias in VOI contouring. Standardised uptake values were evaluated for each tumour VOI with (SUL) and without (SUV) lean body mass correction. The calculation methods used within the HybridViewer Hermes software are defined as follows

$$SUV = A * (W/D) * 1000g$$

$$SUL = A * (LBM/D) * 1000g$$

where:

A= Activity Concentration in Bq/cc (calibrated pixel value)
W=Patient weight in Kg
H=Patient height in cm
D=Injected dose in Bq (decay corrected injected dose to time of activity concentration measurement)
LBM= Lean Body Mass in kg; calculated using the formula $1.07*W=148*(W/H^2)$

The mean and maximum and peak values were recorded for each VOI isocontour structure. Mean was defined as the average value of voxels and maximum as the highest voxel value within the VOI structure. Peak was defined according to PERCIST criteria as the mean value of voxels in the hottest 1 cm^2 constrained inside the VOI structure (62). The most widely cited parameters in the literature (SUVmax, mean and peak) and the PERCIST recommended SULpeak were selected for reporting within in the eCRF.

Exploratory application of PERCIST recommendations for FDG image evaluation was performed (62). To minimise radiation exposure to study participants, the FDG static scan was obtained using a maximum of two bed positions located over the breast and nodal regions and did not encompass sufficient liver for background estimation. Therefore reference background activity was derived on each patient scan using the descending aorta blood pool VOI, defined as a 1 cm diameter cylinder extending over 2 cm in the Z axis within the descending aorta at a site away from diaphragmatic motion artefacts (62). SULmean and s.d. for each VOI were recorded and used to calculate background activity according to the formula

Background = 2(blood pool SULmean + s.d.)

PMod Biomedical Image Quantification software package version 3.4 (PMOD Technologies Ltd, Zurich) was used for quantitative dynamic scan evaluation. The anonymised Dicom PET and CT image data was imported from the HERMES workstation and decay corrected within PMOD where this had not been performed on the scanner. Tumour volumes of interest (VOI) were defined using the 40% isocontour within a manually drawn tumour constraint on a suitable late time point frame. The arterial plasma input function was generated from VOIs defined on a suitable early frame by placing a 1cm diameter cylindrical volume extending over 2 cm in the Z axis within the descending thoracic aorta. VOI and pharmacokinetic analyses were performed using the PVIEW PKIN tools of the PMOD Biomedical Image Quantification software package (version 3.4, PMOD Technologies LTD, Zurich, Switzerland). For FDG scans, the overall influx rate constant (K_i) in min^{-1} and the metabolic rate of glucose (MRGlu) in $mmol \cdot l^{-1} \cdot min^{-1}$, which equals K_i times blood glucose concentration were calculated for each breast tumour VOI using non-compartmental (Patlak graphical analysis) and compartmental modeling (FDG two-tissue compartment model, FDG 2TCM). In patients with measured blood activity derived from a single venous sample acquired at 60 minutes post tracer injection kinetic parameters were also calculated by using the manual blood sample to rescale the descending aorta plasma input function..

6.6 Research Tissue Evaluations

Research tumour tissue was obtained pre and post treatment exposure for correlation with change on treatment SUV/SUL (Part A and B) and early MRI response (Part B only). Material included surplus tissue from the diagnostic core biopsy, ultrasound guided pre-chemotherapy research core, ultrasound guided research core biopsy performed following the day 17 PET-CT scan and a research core through the definitive surgical resection specimen. Patients who gave consent had an optional research blood sample (20-50 mls) processed into serum, plasma and DNA taken contemporaneously with routine pre-chemotherapy blood tests or at a later time point according to patient preference. All blood and biopsy material acquired for the purposes of the study is being stored in the Human Tissue Authority (HTA) licensed Guys' and St Thomas' Breast Tissue and Data Bank. The tissue analysis will be completed in two batches by Dr Patrycja Gazinska (PG) and Professor Sarah Pinder (SP).

The planned tissue analysis within the study comprise immunohistochemical assessment of apoptosis (activated caspase 3), proliferation (MIB-1, the S-phase specific replication marker -

geminin, MCM replication fork licensing factors) and glucose uptake biomarkers (GLUT-1 receptor). However to permit future work using the tissue resource generated through study recruitment documentation was prepared to permit future tumour and blood evaluations. Consequently whole tumour genome sequencing techniques and use of specific tests to detect copy number change (e.g neu/Her2), mutations (e.g. EGFR and other protein markers), DNA methylation status, HER2:HER3 dimer (FRET Efficiency) changes in expression of known and as yet unknown oncogenes, tumour suppressor genes and mRNA expression markers of proliferation genes, and analysis of blood derivatives for DNA, RNA and protein markers of the cancer biology are permitted within the protocol and have regulatory approval.

6.7 Statistical Plan

Trial statistical support was provided by Dr Lucy Kilburn and Ms. Holly Tovey (Institute of Cancer Research Clinical Trials and Statistics Unit, ICR-CTSU). The design of the statistical analysis plan sought to ensure bio-statistical validity and applicability of the generated data to future multicentre PET response validation studies within the dual constraints of comparative rarity of the TNBC subset and cost which both limit numbers possible in an exploratory PET feasibility study.

6.7.1 Sample Size and power

The primary objective of TNPET-01 was to assess whether early PET imaging using either tracer was associated with standard MRI response. For the tracer progressing to part B, correlation of ≥85% between SUV change and % change in the sum of the longest diameter of the breast cancer lesions using RECIST measurements on breast MRI would indicate SUV response to be a good guide to standard MRI response (at 3 cycles). A correlation of 60% or less would be considered unreliable. To achieve this required a total of 20 patients for the tracer progressing to part B, using a one-sided alpha of 10% and a power of 85%.

Based on comparable data in non-breast malignancies (147) and feasibility considerations, Part A repeatability assessment was planned with 10 patients (5 per tracer) prior to selection of a single tracer for progression to part B unless neither tracer met the predefined criteria in which case the study would terminate. Part A required tracer repeatability measurements to be within ±15% for at least 1 SUV/SUL parameter and SUV/SUL reduction to be at least 20% in 50% of MRI defined responders to be considered for progression to Part B. In the event that both tracers met these repeatability criteria the protocol stipulated that the best overall performing tracer was selected for part B.

6.7.2 Statistical analysis methods

The Part A tracer repeatability primary end point analyses were conducted by the ICR-CTSU using Stata version 11 according to the pre-specified statistical analysis plan. Briefly, SUVmax, SUVmean and SUVpeak and SULpeak measurements determined on scans deemed evaluable following visual assessment were considered on a patient-by-patient basis and on a lesion-by-time point basis. As SUV is known to have log-normal distribution (148) measurements were log transformed before analysis. Differences between the two baseline scans were computed and the data were displayed using Bland Altman plots, constructed on a lesion-by-time point basis and plotted on a log scale (points within ±0.14 indicating that the baseline scan measurements at a given time point were within 15% of each other). To take into account nesting of greater than one lesion within some patients repeatability coefficients were calculated using a mixed effects model nesting lesion site within patient with the log parameter as the dependent variable and fitting models in Stata to estimate the within lesion standard deviation from the model. Repeatability coefficients (RC) were calculated from the within-subject variability of the log transformed measurements using the formula ±1.96 x $\sqrt{2}$ x within lesion standard deviation and back transformed hence representing one scan measurement as a proportion of the other rather than absolute change. Upper and lower RC within the range 0.87-1.14 indicates SUV parameters met the predefined 15% criteria. For the patient-by-patient analysis, the individual parameters were summarised using the average value across lesions and scan-times from the same patient. For the lesion-by-time point analysis all individual lesion/time points were analysed.

The final Part B primary and secondary endpoint analysis will be conducted by the ICR-CTSU on completion of full study recruitment. Percent change in SUV parameter on the follow-up day 17±3 scan and % change in the longest MRI defined tumour dimension will be assessed for normality and the appropriate correlation coefficient with a 95% confidence interval (Pearson's correlation coefficient or Spearman's Rank correlation if non-parametric) calculated. Secondary endpoints will be analysed using the relevant summary statistic.

Additional exploratory analyses included within this thesis and the preliminary Part B primary and secondary endpoint analysis (chapters 8 and 9) were performed by JG using SPSS version 22 statistical software and Excel using the appropriate summary statistic according to the statistical plan. Continuous variables have been summarised by mean and standard deviations or by median and range according to the skewness of the distribution. Frequencies and percentages have been used to summarise categorical data; paired t-test for parametric comparison of means or Wilcoxon matched-pairs signed-ranks test for non-parametric data comparisons. Correlations were evaluated using Pearson's correlation coefficient where the

assumption of normality holds. The exploratory repeatability analyses provided in the results chapters but not performed by the ICR-CTSU as part of the trial statistical analysis plan were calculated using the method described above or for non-nested data by calculating the 95% confidence about the mean difference of the log transformed data (149)

6.8 Supporting work for trial set up

6.8.1 Supporting document preparation

In parallel to study protocol preparation, I designed the patient information sheet and developed other supporting documentation including the clinical components of the Investigator Brochure (IB) and Investigation Medicinal Product Dossier (IMPD) to meet the regulatory requirements for clinical trial of investigation medicinal product (CTIMP) classification.

Lay opinion with regard to patient acceptability of the proposed study protocol was sought from representatives of the GSTT patient consumer group comprised of current and historically treated cancer patients. The committee strongly endorsed the study design incorporating novel imaging and serial research tissue. This feedback and their specific recognition of the importance of strategies to improve response evaluation from a patient perspective and input into the patient information sheet facilitated progress through the ethical review process, REC reference 11/L0/1492.

A requirement of all Kings Health Partners/Guys and St Thomas' NHS Foundation Trust (KHP-GSTFT) sponsored CTIMPs is that data is captured within an eCRF. I produced design specifications for each component of the eCRF content and worked closely with the responsible programmer within the Kings Health Partners Clinical Trials Office (KHP-CTO) programmer to ensure consistency of the created data fields with source medical record data and tested the database to ensure fitness for purpose prior to going live. KHP-CTO data extraction for formal analysis by the ICR-CTSU was planned at scheduled analysis time points (end of part A and end of part B).

6.8.2 Neoadjuvant chemotherapy sequencing for study participants

Future neoadjuvant trials in TNBC are anticipated to address questions of efficacy comparing novel agents with the taxane component of standard care. At the time of this protocol design, standard sequencing comprised initial anthracycline component regardless of breast cancer phenotype. Review of clinical data informing sequence in both neoadjuvant/adjuvant settings

indicated enhanced dose delivery, reduced toxicity and pathological response advantage (19, 150-153) with initial taxane scheduling and beneficial reduction in anthracycline cross-resistance compared to the reverse in the preclinical setting (154). Furthermore recent national and international trial protocols (e.g. neo ALLTO, NSABP-40, I-Spy 2 and SOLD, the control arm of the currently recruiting UK randomised phase 3 neoadjuvant study ARTemis and the multicentre Neo-tAngo trial) have selected taxane first sequencing in the standard of care arms.

Subsequent to data review and with the approval of the breast oncology consultant body I updated the South East London Cancer Network (SELCN) systemic guidance to permit taxane first sequencing in patients with HER2-ve breast cancer receiving treatment in the neoadjuvant setting. This change to the SELCN protocols was formalised following approval by the breast tumour working group (Jan 2012), revised prescription proformas approved by pharmacy (March 2012) and implemented on electronic (CIS) prescribing platform (May 2012). The delivery of taxane first chemotherapy sequencing as standard to study participants ensures a dataset reporting PET-CT efficacy evaluation in TNBC following uniform therapeutic exposure.

6.8.3 Radiology considerations

The importance of support from the radiology team in terms of RECIST reporting and acquisition of research biopsy tissue was recognised as a critical element for the success of the study early in the protocol development. To ensure adequate resources for their support of this and related studies incorporating tissue biomarker endpoints, I successfully sought funding to allow purchase of 4 research bio-specimen procurement containers (for storage of research tissues in the radiology department and in transit to the Breast Tissue Bank) and towards a sonographer advanced practitioner from the London South Clinical Research Network.

In an effort to develop a non-invasive measure of pCR there is interest in MRI parameters that have the potential to predict later response (155). Amongst these, diffusion weighted (DW) MRI is an imaging technique where tissue water diffusivity is measured and quantified as the Apparent Diffusion Coefficient (ADC) (156). ADC is lower in tumours than normal tissues reflecting greater cell density and restriction of water diffusion (157). ADC evaluation increases the diagnostic specificity of MRI at staging of breast cancer (158) and quantifiably increases in responding disease (157, 159) without compromising RECIST evaluation (160-162). Consequently DW-MRI is routinely included in breast assessment at many centres including Guy's and St Thomas'. There is emerging evidence that therapy induced increases in ADC predate MRI detectable size change (163) and may be informative from as early as 1 cycle. Recognising the potential utility for early MRI with ADC evaluation, the study protocol includes

an additional research MRI scan after one cycle of chemotherapy in part B. This provides a unique opportunity to investigate exploratory performance estimates of early ADC which can be compared with on treatment PET assessment of proliferation or glucose metabolism, the RECIST response and definitive RCB response providing a stronger rationale for incorporation of functional imaging in future studies. High temporal resolution DCE MRI protocols evaluate tumour microvasculature by permitting pharmacokinetic analysis but currently do not have the capability to encompass the whole breast with sufficient spatial resolution to fulfil clinical requirements for reporting using the BIRAD lexicon and RECIST response monitoring as well (21). This and the additional requirement for a reproducibility DCE MRI at time of the first pre-treatment MRI to validate the technique mean that exploratory evaluation using DCE MRI sequences within this study would mandate further patient attendances for research MRI scans at standard and post cycle 1 time points (i.e. a total of 8 MRI scans on separate visits) to prevent compromise of clinical scans and evaluation of the PET primary endpoint. In view of the additional costs and threat to patient recruitment inherent to the addition of research DCE MRIs were not felt to be justified in the TNPET01 PET feasibility study. Consequently DCE sequences were not been included in this exploratory study where the primary research focus was feasibility of early PET-CT to determine later MRI RECIST response.

6.8.4 Funding

I successfully secured funding from the Comprehensive Cancer Imaging Centre CCIC (Part A) and the Guys and St Thomas' Charity (Part B). A prior invited full application to CRUK BIDD (November 2011) received supportive reviewers comments but was unsuccessful in its funding application.

6.9 Achieved trial timelines for study set up and opening

The timeline for completion of the required steps to first patient recruitment is summarised in figure 4.2. Non-substantial amendments were required to encompass MHRA mandated toxicity assessments (telephone call at 24 hours post FLT PET scan) and specifying that randomisation would be performed by the ICR-CTSU rather than by the study sponsor as originally planned in the set up phase. The sabbatical taken by the eCRF programmer and the time taken for the service level and co-sponsorship agreements contributed to delays to the set up time-frame.

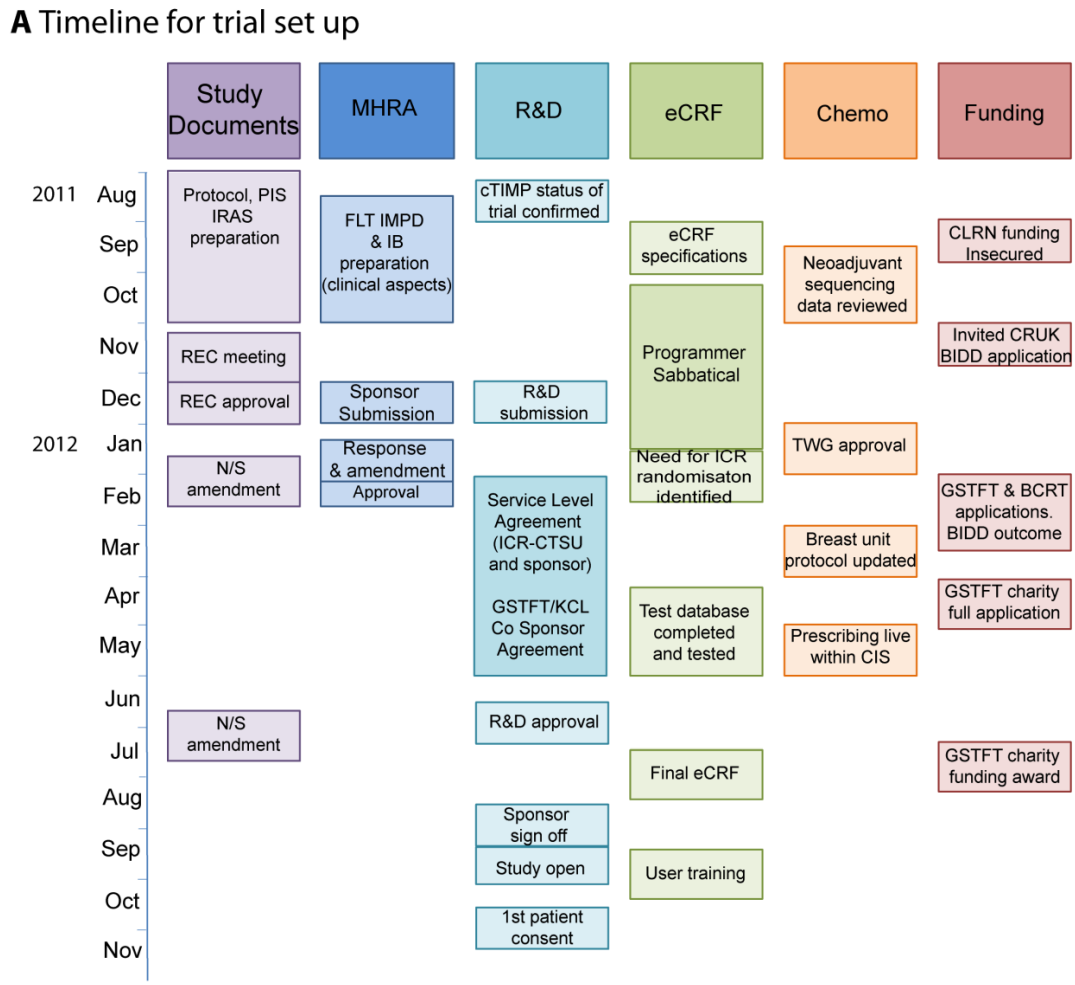


Figure 6-2 Summary of achieved trial time lines in set-up phase

Key: BCRT Breast Cancer Research Trust, PIS Patient information sheet, IRAS Integrated research application service, REC Research Ethics Committee, N/S non substantial, IB Investigator Brochure, IMPD Investigational Medicinal Product Dossier, cTIMP clinical trial of investigation medicinal product, CRUK Cancer Research UK

6.10 Study accrual and protocol amendments

The study opened later than anticipated in August 2012 and the first patient gave consent in November 2012. Based on unit data an accrual rate of at least 12 patients per year was anticipated and the initial statistical plan anticipated full trial accrual of 30 patients to complete in 24-30 months and progression of both tracers to Part B if repeatability criteria were met. Recognising the delays to study opening, inward referral from centres within the South East London Cancer Network (SELCN) was sought (presentation to the study to the SELCN and to the Queen Elizabeth Hospital and Kings College Hospital MDMs). Support was agreed with the relevant surgical teams for inward referral of eligible and potentially interested patients for treatment at GSTT to facilitate accrual. Unfortunately cyclotron failure

(Feb/March 2013), PET centre refurbishment (Sept-December 2013) and FLT production failure (Feb-July 2014) had a significant impact on recruitment incurring a cumulative delay of 12 months such that by July 2014, 9 of 11 recruited patients had completed all Part A evaluations. Recognising the ongoing radio pharmacy issues thwarting reliable FLT production the decision was made to seek REC approval for a substantial protocol amendment permitting repeatability evaluation at 9 complete patient datasets and progression of the best performing single tracer meeting repeatability criteria through to part B rather than both as had been originally planned and approved at study opening. This amendment planned for total number of patients scanned using the single tracer to be increased from 15 to 20 for Part B analysis, thus tightening the confidence interval for the response assessment data Following REC and MHRA approval of this substantial amendment the TSC reviewed the data in October 2014 and the first patient was recruited to part B later that month. At the time of writing (June 2016) 13 of 18 recruited patients have competed all research evaluations and recruitment to Part B continues.

6.11 Conclusions

The aspiration of improving response evaluations strategies in TNBC with a view to tailoring therapeutic decisions to the individual response requires phenotype specific validation of imaging biomarkers with reference to relevant clinical outcomes. The study was designed to address questions relating to tracer selection, test-re-test repeatability and its relationship to therapy induced change for optimisation of imaging response strategy and inform larger scale trial design.

As described in section 6.9 and 6.10 un-anticipated difficulties were encountered in both the set up and recruitment phases. Although recruitment to part A has completed; the tempo has precluded completion of Part B recruitment within the MD timeframe. Steps taken to mitigate the impact of set-up delays and accelerate recruitment through inward referral from centres in the network have been frustrated by direct loss of 12 months accrual opportunities resulting from the PET centre refurbishment and the FLT IMP production failures. The protocol had been originally worded to require replacement of patients in the event of failure of an individual to complete of PET imaging at all-time points with the aim of protecting against study completion with an incomplete dataset and consequent loss of statistical power. Had this replacement requirement not been in place the Part A analyses could potentially have taken place earlier when it became clear that FLT production could not be readily reinstated. With hindsight omitting the patient replacement strategy for Part A from the outset may have been advantageous, but carried the risk of accrual without achieving statistical validity. Recognising

the impact of the unresolvable FLT production failures the July 2014 protocol amendment successfully permitted Part A evaluation at 9 completed imaging datasets and progression of a single tracer to Part B response evaluation reducing the total number of participants from 30 to 25. At the time of writing 13 of 18 recruited patients have competed all research evaluations and recruitment to part B continues. Permission has been obtained from GSTFT charity funders for a time only extension until Feb 2017 to allow adequate time for study completion.

7 Results: SUV Repeatability

7.1 Aims

This chapter presents the first repeatability parameters for static scan acquisitions using the FLT or FDG tracers in the triple negative breast cancer phenotype. As described in chapter 4, participants enrolled in part A underwent two pre chemotherapy PET-CT scans using either the FLT or FDG tracer according to randomisation. Static scan acquisitions were performed at 90, 120 and 180 minutes post tracer in the FDG group and at 90 minutes post injection in the FLT group. SUVmax, SUV mean, SUVpeak and SULpeak were measured for all tumour lesions in each static acquisition for primary part A endpoint repeatability analysis determining tracer progression to part B response evaluation.

Further exploratory analyses presented in this chapter are measures of repeatability in tumour lesions defined using PERCIST, hottest lesion and within breast lesions only evaluability criteria.

7.2 Patient Cohort

A total of 11 patients were randomised within part A. 5 participants scanned using the FDG and 4 participants randomised to FLT completed both baseline scans and were included in the analyses. In the FLT group 1 participant withdrew consent following her first PET scan and a further patient was withdrawn from the study without receiving any allocated research interventions due to FLT production failure.

Table 7.1 summarises baseline patient and tumour characteristics. All tumours were confirmed ER and HER2 negative and were generally high grade no special type (NST). Stated nodal involvement was cytology confirmed. Primary tumours were at least stage T1c by MRI, and with one exception were greater than T2. This smallest tumour had clinical measurements of 13mm by MRI, 17mm by USS and 25mm by palpation, clinical stage T2.

Demographics		FLT (n=4)	FDG (n=5)	Overall (n=9)
Age (years)	Mean (s.d.)	40.3 (10.7)	43.6 (5.0)	42.1 (7.7)
	Median (min-max)	39 (30-53)	44 (36-49)	44 (30-53)
Race	Black	1 (25%)	3 (60%)	4 (44.4%)
	White	3 (75%)	2 (40%)	5 (55.6%)
Body Mass Index	Mean (s.d.)	23.35 (4.52)	29.44 (5.28)	26.69 (5.68)
	Median (min-max)	23.4(17.6-28.6)	27.5 (23.9-37.6)	27.1 (17.6-37.6)
Tumour characteristics				
Primary tumour size	Mean (s.d)	31.0 (5.72)	32.8 (14.5)	32.0 (10.87)
	Median (min-max)	30.5 (28-38)	35 (13-51)	33 (13-51)
Nodal involvement	Positive	1 (25%)	2 (40%)	3 (33.3%)
	Negative	3 (75%)	3 (60%)	6 (66.7%)
Number of lesions	1	3 (75%)	3 (60%)	6 (66.7%)
	2	1 (25%)	2 (40%)	3 (33.3%)
Histological type	NST	2	5	7
	Ductal	2	0	2
Grade	2	1	0	1
	3	3	5	8
ER score	0	4	4	8
	1	0	0	0
	2	0	1	1
HER2 score	0-1	4	4	8
	FISH -ve	0	1	1

Table 7-1 Baseline characteristics by allocated tracer group

Categorical variables are reported by frequencies and column percentages. s.d. = standard deviation. Primary tumour size according to MRI longest dimension, mm. Number of lesions includes breast and nodal involvement where present.

The same scanner was used for PET imaging at baseline and post cycle 1 study visits. The maximum repeatability acquisition interval was 5 days (Table 7.2). Chemotherapy commenced

within 10 days of the first baseline scan and response scans took place in the third week following cycle 1 in all study participants.

	Scan Interval		
	Repeatability scan Interval Median (range) days	1st baseline to 1st chemotherapy Median (range) days	1st chemotherapy to PET response scan Median range) days
FLT group (n=4)	3.5 (3-5)	6.5 (5-7)	19 (17-20)
FDG group (n=5)	4 (3-4)	8 (6-10)	18 (17-18)

Table 7-2 PET visit interval

Interval between baseline repeatability scans, first baseline and first chemotherapy cycle and the post cycle 1 PET response. Median and range (days) are reported.

7.3 Static scan acquisition parameters

FDG scans were acquired after minimum 6 hours fast time. All follow up scans commenced within 5 minutes of the scan 1 acquisition time (Figure 7.1 A). No patient had serum glucose level >6.2 (normal range 4.4 to 6.1mmol/L). The group mean serum glucose levels were 5.04 mmol/L (sd 0.73), 5.42 mmol/L (s.d 0.51) and 4.82 (s.d 0.69) at the first baseline, second baseline and post cycle 1 visits respectively. Glucose levels did not statistically differ between baseline visits (p= 0.31) or post cycle 1 visits (p=0.19). The mean injected activity was 188.4 (s.d 7.9) MBq, 183.2 (s.d 5.6) MBq and 188.5 (s.d 6.6) MBq for FDG scans performed at the first baseline, second baseline and post cycle 1 visits respectively. Injected activity did not statistically differ between the two baseline visits (p=0.38) or post cycle 1 (p= 0.98) and was within ±10% for each individual (range -8.20 to 7.70%).

All FLT follow up scans commenced within 12 minutes of scan 1 acquisition time (Figure 7.1B). The FLT mean injected activity was 187.6(s.d 6.6) MBq, 185.2 (s.d 9.5) MBq and 183.7 (s.d 6.2) MBq at the first baseline, second baseline and post cycle 1 visits respectively. Injected activity did not statistically differ between the two baseline visits (p=0.767) or post cycle 1 (p=0.213) and was within ±12% for each patient (range -6.84 to 11.23%).

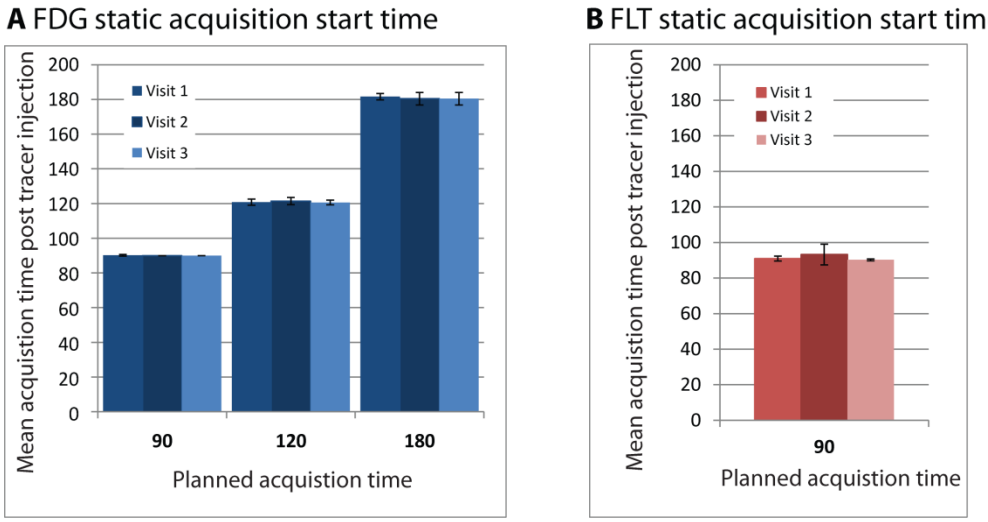


Figure 7-1 Static scan acquisition time for baseline and follow-up scans

A Follow up acquisition time for static scans scheduled to commence at 90, 120 and 180 minutes following FDG tracer injection. Mean and standard deviation (s.d) shown for each time point. No statistical difference in acquisition time was present for any follow up scan comparison (p>0.3 for all paired comparisons). B. Follow up acquisition time for static scans scheduled to commence at 90 minutes following FLT tracer injection. Mean and s.d shown for each 90 minute scan acquisition. No statistical difference in acquisition time was present for any follow up scan comparison (p>0.4 for all paired comparisons).

7.3.1 Lesion assessment on baseline repeatability scans

All scans were deemed suitable for analysis following visual inspection by an experienced nuclear medicine physician (MD or SB). The breast primaries were visualised in all patients. Axillary lesions were visualised in 3 patients (1 from FLT group and 2 from FDG group). Tables 7.3 and 7.4 summarise the SUV parameters for the FLT and FDG tracer groups respectively. SUVmax and SUVmean standard uptake parameters were evaluable across all lesions but peak parameters could not be derived within HERMES for one (axillary) lesion in the FLT group. Using the 40% isocontour In the FDG group, peak parameters of breast lesions could not be derived on all scans at all-time points in two patients (MRI breast lesion size 40mm and 13mm) and in the axilla in 1 patient.

		Scan1	Scan 2
SUVmax	n	5	5
	Mean (s.d.)	6.32 (2.63)	6.84 (2.83)
	Min-max	2.9-9.8	2.8-10.6
SUVmean	n	5	5
	Mean (s.d.)	3.93 (1.73)	4.29 (2.02)
	Min-max	1.8-6.3	1.8-7.1
SUVpeak	n	4	4
	Mean (s.d.)	5.25 (2.75)	5.31(2.79)
	Min-max	2.4-8.8	2.5-9.1
SULpeak	n	4	4
	Mean (s.d.)	3.76 (1.78)	3.80 (1.74)
	Min-max	1.7-5.7	1.8-5.9

Table 7-3 Summary of standard uptake parameters for the FLT tracer

Results are reported for the two repeatability scans. Standard uptake values, mean, standard deviation (s.d.) and range shown for each parameter. SUVpeak and SUVmax could not be derived for the axillary lesion in one patient.

		Scan 1			Scan 2		
		90 min	120 min	180 min	90 min	120 min	180 min
SUVmax	n	7	7	7	7	7	7
	Mean (s.d.)	11.18 (8.02)	12.36 (8.95)	12.75 (10.81)	11.02 (7.33)	11.52 (7.43)	12.98 (8.91)
	Min-max	5.0-28.0	4.9-30.9	6.0-36.8	4.6-25.9	5.7-26.8	5.8-31.4
SUVmean	n	7	7	7	7	7	7
	Mean (s.d.)	6.91 (5.29)	7.66 (5.96)	8.83 (7.42)	6.77 (4.83)	7.15 (5.02)	8.05 (5.94)
	Min-max	2.9-18.2	2.9-20.2	3.6-24.8	2.8-16.8	3.2-17.8	5.8-31.4
SUVpeak	n	5	4	5	4	4	5
	Mean (s.d.)	11.09 (8.16)	13.21 (10.82)	14.06 (11.12)	11.53 (8.78)	11.90 (8.46)	12.69 (8.47)
	Min-max	5.5-25.3	6.1-29.2	6.6-33.5	5.7-24.4	6.2-24.2	6.7-27.4
SULpeak	n	5	4	5	4	4	5
	Mean (s.d.)	7.35 (5.51)	8.89 (7.24)	9.28 (7.48)	7.78 (5.89)	8.03 (5.69)	8.38 (5.71)
	Min-Max	3.7-16.9	4.2-19.6	3.3-22.4	3.6-16.3	4.1-16.2	4.5-18.3

Table 7-4 Summary of standard uptake parameters for the FDG tracer

Results are reported for the two repeatability scans at the 90, 120 and 180 minute static acquisition time points. Standard uptake values, mean, standard deviation (s.d.) and range shown for each parameter. SUVpeak and SULpeak could not be derived for one axillary lesion at any acquisition point. For breast lesions the peak parameters could not be fully derived at 90 minutes (1 patient scan 1, 2 patients scan 2), 120 minutes (2 patients scan 1 and 2) and 180 minutes (1 patient).

7.4 Per Protocol Repeatability results

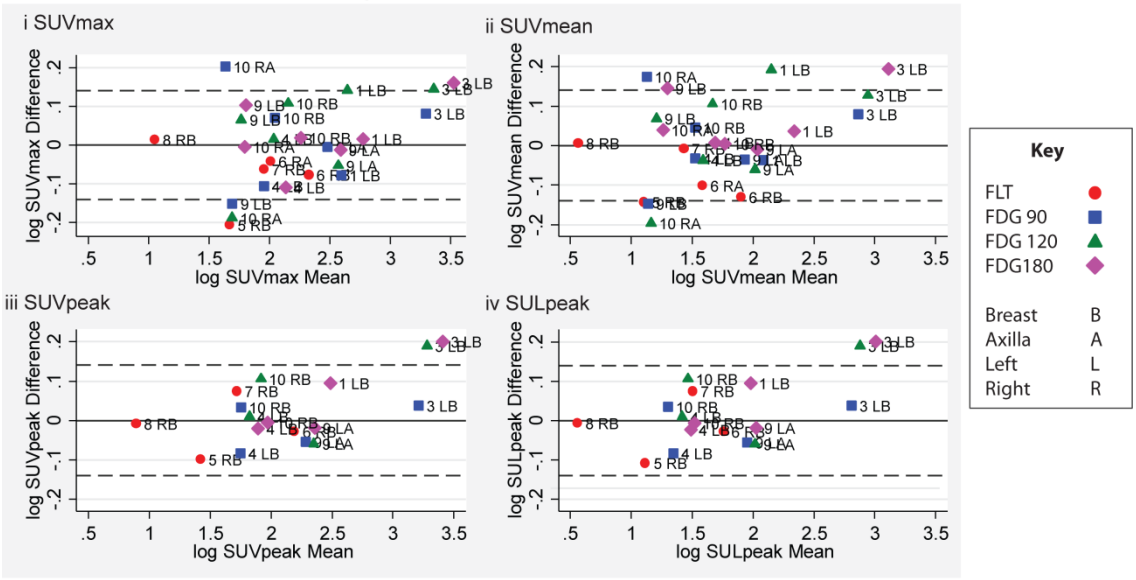
For Part A repeatability end point analysis Bland-Altman plots were constructed for each tracer on a lesion by time point basis (Figure 7.2 A). For the majority of lesions the % change between scans is within 15%. Both tracers present lesions with differences in some SUV parameters greater than 15% however for all lesions and timepoints the SUV parameters did not differ by more than 22% on the two scans (±0.2 on the logarithmic scale). Greater variability was present for SUVmean and max parameters than for either SUV or SULpeak parameters. No

repeatability benefit for lean body mass corrected Peak is evident for either tracer or at any timepoint.

Figure 7.2B reports the repeatability coefficents (RC) of the standardised uptake parameters for each tracer on a patient by patient basis, expressed as a ratio of scan 1 to scan 2. The repeatability of FDG is within 15% for SUVmean and SUVmax and within 16% for SUVpeak and SULpeak. The FLT repeatability is within 15% for SUVpeak and SULpeak, but higher for SUVmax and SUVmean (within 25% and 20% respectively). Therefore the per-patient repeatability criteria (specified in as primary analysis in the Part A analysis plan) were fulfilled by both tracers for at least one SUV measure.

As a sensitivity analysis RC were also computed on a lesion-by–time point analysis to allow for the nested nature of the data (patients with more than 1 lesion) (Figure 7.2C). At the 90 minute time point FDG lesion repeatability for SUVpeak and SULpeak is within 12% increasing to 24% for SUVmax. Lesion repeatability is higher for other parameters and time points; up to 28% for SUVmean at 120 minutes. For FLT, lesions-by-time point RC's are generally higher than on per patient analysis indicating up to 23% change between the two scans for SUVmax.

A Bland-Altman Plots for SUV parameters



B Repeatability coefficients per tracer type (patient by patient analysis)

	Repeatability	
	FLT (n=4)	FDG (n=5)
SUVmax	0.80-1.25	0.87-1.15
SUVmean	0.83-1.20	0.87-1.15
SUVpeak	0.88-1.13	0.86-1.16
SULpeak	0.88-1.14	0.86-1.16

C Repeatability coefficients per tracer type (lesion-by-time point analysis)

	FLT	FDG		
	90 minutes	**90 minutes**	**120 minutes**	**180 minutes**
SUVmax	0.82-1.23	0.80-1.24	0.80-1.26	0.85-1.18
SUVmean	0.83-1.21	0.83-1.21	0.78-1.28	0.83-1.20
SUVpeak	0.88-1.13	0.90-1.12	0.80-1.25	0.82-1.22
SULpeak	0.88-1.14	0.90-1.12	0.80-1.25	0.82-1.22

Figure 7-2 Part A repeatability evaluation

A. Bland Altman plots of log transformed SUV measurement for the two baseline repeatability scans. Plots per lesion, time point and allocated tracer for i log SUVmax, ii log SUVmean, iii log SUVpeak and iv log SULpeak. Points on the graph are labelled with the individual number and lesion location. Lesions within -0.14 and 0.14 (dashed lines) indicate baseline scan measurements were within 15% of each other. B. Repeatability coefficients (RC) for SUV parameters and each tracer on a patient by patient basis. RC expressed as ratios of the two scans, 0.87-1.15 and 0.83-1.20 representing lower and upper bounds of 15% and 20% change between scans, SUVmean and SUVmax (FDG) and SUVpeak and SULpeak (FLT) parameters are within 15%. C. RC for SUV parameters on a lesion-by-time point analysis

7.5 Exploratory lesion repeatability results

Lesion repeatability was considered using the subset of lesions evaluable according to criteria adopted in published breast cancer FDG response evaluation studies, namely breast lesions only; the hottest single lesion on the baseline scan (56), and those meeting PERCIST criteria with evaluable Peak lesions (62). Using hottest lesion criteria the primary breast tumour was target for response evaluation in 4 of the 5 patients (Figure 7.3 A). SUVpeak could not be be determined on the second baseline acquisition within the 40% isocontour. Thus using PERCIST criteria Patient 1 was not evaluable. In this patient the primary breast tumour measured 40mm on the baseline MRI and baseline SUVmax values were 12.88 and 13.93 on the first and second 90 minute static scan acquisitions respectively.

At the 90 minute time point SUVmax and SUVmean repeatability was within 12%, (corresponding to difference 0.11 on the log scale) using the hottest lesion and PERCIST lesion selection method, (Figure 7-3B). Repeatability for peak parameters was within 15% regardless of methodology but only 4 of 5 FDG patient scans were evaluable using PERCIST criteria (Figure 7-3B). At the 120 minute time point using the hottest lesion and PERCIST criteria similar tightening of baseline repeatability was present, to within 16%, corresponding to difference 0.15 on the log scale for SUVmax and SUVmean parameters. At 180 minutes repeatability was within 12% for SUVmax, 21% for SUVmean and 22% for the two peak parameters.

A 90 minute FDG lesion evaluability comparison

Patient (FDG)	Lesion Location	Lesion Evaluability Method (90 minute FDG scan)			
		Visual Inspection	Breast	PERCIST	Hottest Lesion
1	Breast	Yes	Yes	No	Yes
3	Breast	Yes	Yes	Yes	Yes
4	Breast	Yes	Yes	Yes	Yes
9	Breast	Yes	Yes	No	No
9	Axilla	Yes	No	Yes	Yes
10	Breast	Yes	Yes	Yes	Yes
10	Axilla	Yes	No	No	No

B Bland Altman plots according to lesion evaluability criteria

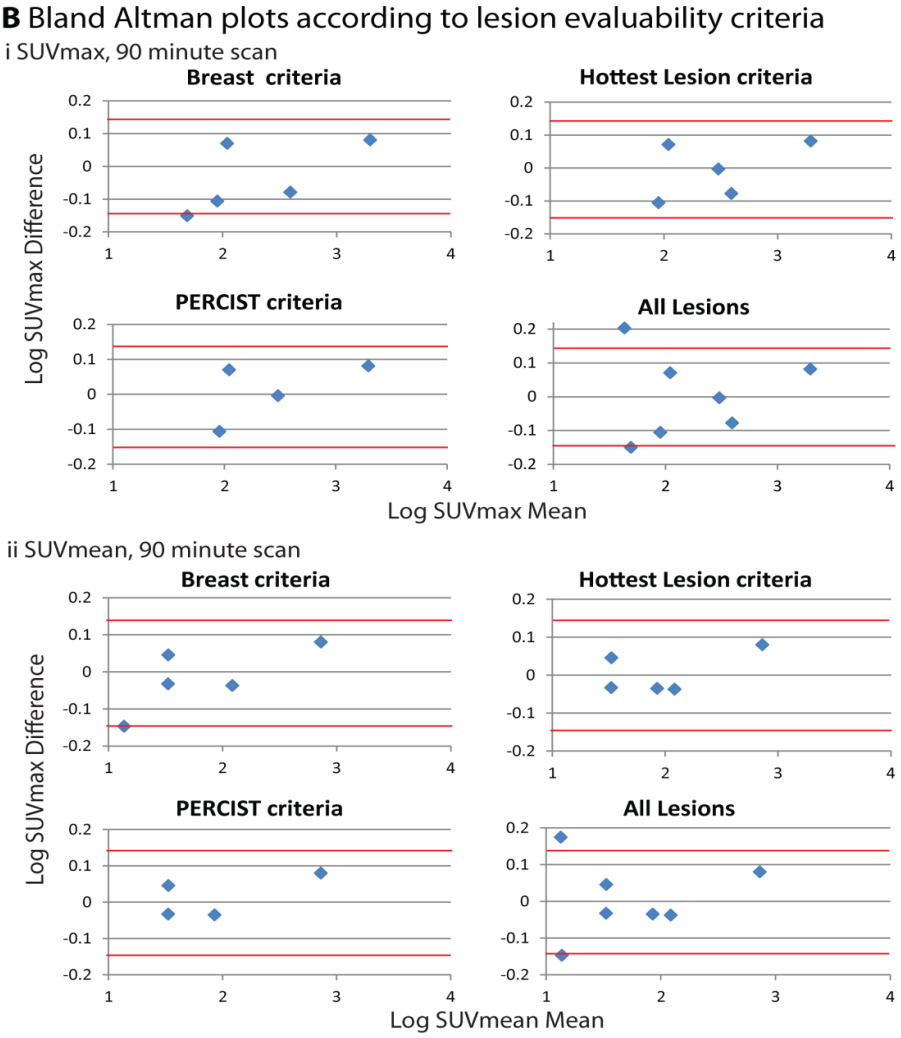


Figure 7-3 Baseline tumour repeatability considered in relation to methodology for defining response

A Lesion evaluability at the 90 minute time point considered according to methodologies adopted in published response evaluation studies. The total number of baseline lesions suitable for response analysis varies from 7 (visual inspection method) to 4 (PERCIST).

B. Bland Altman plots of log transformed measurements for the two baseline repeatability scans. Plots per lesion, 90 minute time point for i log SUVmax, ii log SUVmean according to methodology used to define lesion evaluability. Lesions within -0.14 and 0.14 (dashed lines) indicate baseline scan measurements were within 15% of each other.

7.6 Impact of scan reconstruction method on repeatability

All FLT PET-CT scans were performed using the GE Discovery VCT 64 slice PET-CT scanner however PET centre refurbishment necessitated scanner transition to the GE Discovery 710 64 slice PET-CT scanner after the first 3 FDG imaged participants. Recognising the new scanner capability and published data demonstrating improved visualisation of small lesions with time of flight reconstruction (TOF), all clinical and research imaging following scanner installation was performed using TOF and 256 matrix reconstruction in contrast to the 128 matrix used previously (164-166). To exclude the possibility of detrimental impact of this reconstruction technique on SUV repeatability within the TNPET study, the images acquired on the GE Discovery 710 64 slice PET-CT scanner were reconstructed using 128 matrix and 256 VPHD matrices in addition to TOF for visual comparison of the 3 sets of log transformed SUVmax and SUVmean parameters using Bland Altman plots.

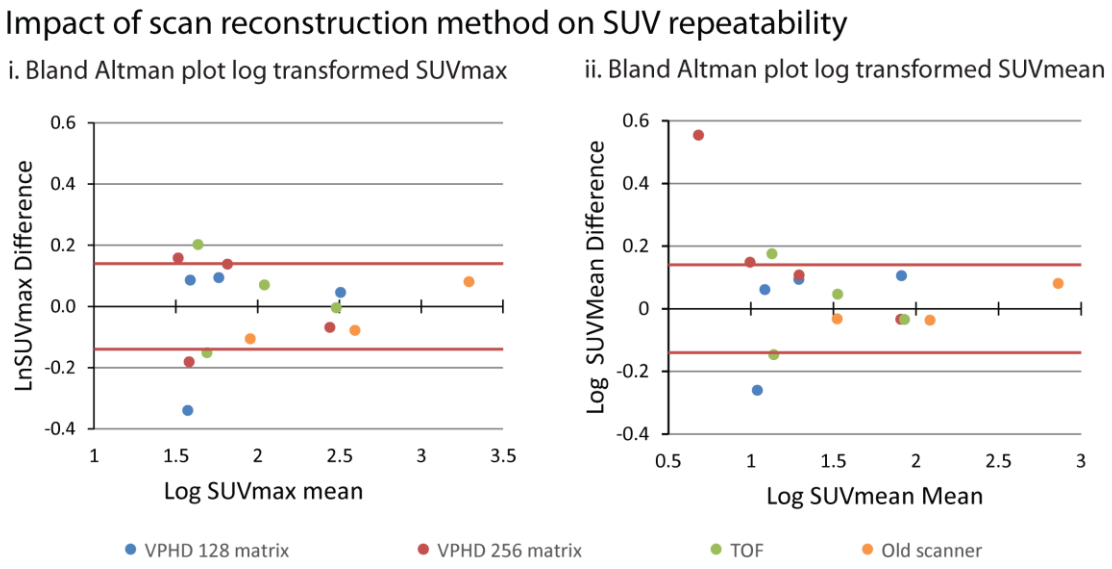


Figure 7-4 Comparison of scan reconstuction method (90 minute scan acquistion)

Bland Altman plots of log transformed 90 minute measurements for the two baseline repeatability scans for images acquired post PET centre refurbishment (2 participants, 2 lesions per participant). Plots per lesion, for the three scan reconstruction methods, 128 matrix (blue), 256 matrix (red) and time of flight (TOF, green). Values for lesions imaged prior to PET centre refurbishment are shown for comparison (orange). i log SUVmax, ii log SUVmean. Lesions within -0.14 and 0.14 (dashed lines) indicate baseline scan measurements were within 15% of each other. Irrespective of reconstruction technique all hottest lesion targets were within +/- 0.14 bounds.

Considering all lesions, SUVs derived using the TOF reconstruction exhibited the narrowest lesion difference between the two baseline 90 minute scan acquisitions (Figure 7-4). For Log SUVmax the mean difference and 95% confidence interval for the difference were -0.03 (0.84 to 1.36); -0.08 (0.77 to 1.39) and -0.002 (0.78 to 1.27) for 128, 256 and TOF reconstructions respectively. For LogSUVmean the mean difference and 95% confidence intervals for the difference were 0.012 (0.77 to 1.29); 0.11 (0.66 to 1.51) and 0.007 (0.98 to 1.01) for 128, 256 and TOF reconstructions respectively. Hottest lesion targets were within the protocol defined 15% repeatability bounds using all three methods. Similar repeatability performance of TOF relative to the other reconstructions was present at the 120 and 180 minute time points.

7.7 Background uptake

The mean and SD of the calculated background for the baseline and response FDG PET acquisitions are shown in Table 7.5. No statistical difference in the calculated blood background on the two baseline scans was demonstrated at any static time point (p=0.58, p=0.73 and p=0.74 for 90, 120 and 180 minutes respectively, paired sample t-test). Comparison of group background SUL at first baseline with the post cycle 1 response scan showed significant difference only at the 120 minutes time point (p=0.72, p=0.04, p=0.256 for 90, 120 and 180 minutes respectively, paired sample T-test).

Scan acquisition	**Group background SUL**		
	Baseline scan 1 Mean (s.d)	**Baseline scan 2** Mean (s.d)	**Response scan** Mean (s.d)
90 minutes	2.54 (0.56)	2.68 (0.40)	2.60 (0.42)
120 minutes	2.10 (0.39)	2.17 (0.52)	2.41 (0.48)
180 minutes	2.17 (0.53)	2.12 (0.49)	1.93 (0.18)

Table 7-5 Group background SUL for FDG part A scans

Background SUL calculated for each FDG static acquisition according to the method defined in PERCIST. Mean and standard deviation shown for the group at each of 90, 120 and 180 PET image acquisitions.

All scans with tumour peak parameters evaluable within HERMES met PERCIST criteria for baseline lesion evaluability (SULpeak greater than the calculated background). For follow up response assessment PERCIST recommends background is within 20% of baseline for that

individual. Although the group background was within 20% of group baseline at all time points this criterion was not met when considered on a per patient basis for images acquired at 120 and 180 minutes.

7.8 Response assessment

For tracers to be considered for progression to Part B an SUV/SUL reduction of at least 20% in 50% of MRI defined responders was required in addition to repeatability within ±15% for at least 1 SUV/SUL parameter. All patients with an objective response (complete or partial) or stable disease had drop of ≥20% in SUV and SUL. The single patient with progressive disease (FLT group) had an SUV and SUL drop of less than 20%. Therefore both tracers met the predefined criteria. The % change for each SUV parameter at the post cycle 1 scan compared to the mid-MRI response is shown in Figure 7-5.

In the FDG group the mid MRI reported disease response as complete (1 patient), partial (2 patients) and stable disease (2 patients). Compared to the second baseline scan all patients exhibited change in the peak parameters which exceed the PERCIST suggested threshold change of 30% to discriminate responders at the 90 minute scan acquisition. Similarly the minimum observed drop in SUVpeak was 41.6% and 42% respectively at the exploratory 120 and 180 minutes scan acquisitions. This individual with the poorest level of SUV response had stable disease on her post cycle 3 scan but clinical concern regarding progression necessitated surgery prior to transition to the second component of neoadjuvant treatment. It is therefore possible that the recommended threshold change of 15% (EORTC) and 30% (PERCIST) to define responders is insufficiently stringent in TNBC. However a larger cohort would be required to confirm this observation.

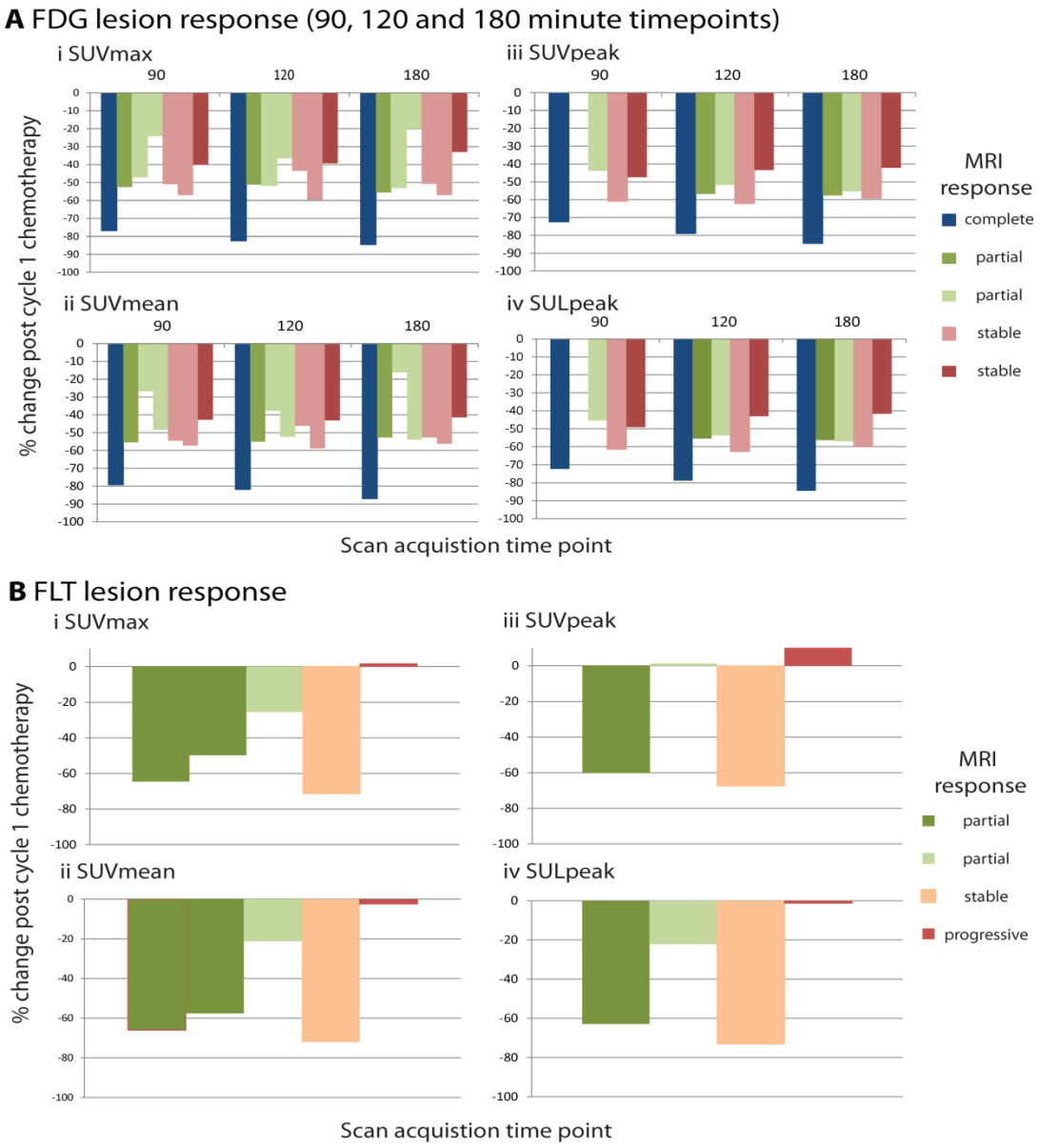


Figure 7-5 Part A PET response as a function of post cycle 3 MRI RECIST response

A. *FDG response at each of 90, 120 and 180 minute acquisitions for i SUVmax, ii SUVpeak, iii SUVmean and iv SULpeak parameters. Colour scale indicates post cycle 3 reference MRI response for each individual. Where 2 bars are present, this indicates evaluation of breast and nodal lesions, in each case the bar with the greatest magnitude of change represents the hottest lesion. The patient denoted in dark red underwent surgery prior to transition to anthracycline chemotherapy due to unequivocal clinical progression despite the MRI findings.*

B. *FLT response for each of i SUVmax, ii SUVpeak, iii SUVmean and iv SULpeak parameters. Colour scale indicates post cycle 3 reference MRI response for each individual. Where 2 bars are present, this indicates evaluation of breast and nodal lesions, the bar with the greatest magnitude of change represents the hottest lesion.*

7.9 Conclusions

PERCIST guidance makes a number of recommendations pertaining to FDG imaging acquisition for serial imaging which include ensuring that follow up scanning occurs within 10 minutes of planned acquisition time post tracer, consistency of scanner and injected dose within +/- 20%. In the FDG repeatability cohort these criteria were met. No comparable recommendations have been published for non-FDG tracer but the FDG acquisition standards were also applied to patients imaged using the FLT tracer. For both tracers image acquisition was optimised with PET imaging performed at fixed intervals after tracer injection avoiding potential confounding impact of technical variability between scans within the same individual on repeatability evaluation (79).

We report the first breast cancer subtype specific repeatability data for commonly reported standardised uptake parameters (SUVmax, SUVmean, SUVpeak and SULpeak) assessed at conventional (90 minutes) and exploratory (120 and 180 minute) acquisition time points. When considering all breast and nodal lesions our data demonstrates that SUV intrinsic variability is 12-24% in both tracers, that is dependent on scan acquisition time and SUV parameter. This intrinsic variability raises questions about the validity of suggested EORTC threshold SUV change of 15% for defining responders in treatment assessment of TNBC. Furthermore this phenotype specific repeatability and preliminary response data indicate SUV threshold change in TNBC will exceed currently recommended 15-30% SUV change for solid tumour chemotherapy response prediction. In the context of an aspiration to ultimately improve outcomes for those with RCB2/3 response by implementing therapy change early in the course of neoadjuvant more stringent criteria may be required for defining PET response as a predictive biomarker for chemo sensitivity. Furthermore although overall repeatability was less robust for SUVmax or mean parameters, where a response evaluation approach considering only the single hottest lesion is adopted repeatability improves to within 12%.

The current data identifies the PERCIST recommended peak parameter (with or without BSA correction) to be problematic in assessment of some breast cancers or involved nodal lesions. Concerns regarding impact of change in weight and BMI through a course of cytotoxic therapy on tracer kinetics and SUV evaluation underlie PERCIST recommendations for lean body mass correction. However unlike the metastatic disease context, in those undergoing neoadjuvant treatment significant cachexia at diagnosis or change in BMI early in the course of chemotherapy would be rare and as expected was not evident in the repeatability population and may explain why LBM correction did not improve either peak evaluability or repeatability. All patients in the test re-test cohort had T1c tumours as measured by baseline MRI and were >2cm by palpation with PET images considered evaluable by experienced nuclear medicine

physicians. Published data evaluating the diagnostic role of PET in breast cancer reports 98% tumours >15mm tumour to be evaluable using SUVmax and recommends minimum stage T1C to be suitable for response monitoring (167). Nevertheless it is possible that limiting eligibility to participants with larger primary breast tumours, for example >3cm, may have improved both repeatability and lesion evaluability according to PERCIST definition of peak. However where this approach has been previously adopted only the SUVmax parameter has been reported (72, 73). As neoadjuvant sequencing is increasingly recommended to patients with TNBC at the lower end of the T2 or T1c stage, particularly in the context of cytological evidence of nodal involvement, the ability to evaluate these smaller lesions is a key requirement of imaging biomarkers for response assessment. Our data suggests advantages in terms of lesion evaluability using SUVmax or mean parameters compared to the peak parameter with or without LBM correction in the neoadjuvant breast cancer setting.

Based on the Part A results and also recognising the problems with FLT reliability, the FDG tracer alone was selected for progression through to part B. Full study completion (anticipated 2017) will inform future use of early FDG-PET as an exploratory biomarker novel therapy neo-adjuvant trials in TNBC. The optimal strategy for evaluating PET response remains undefined but the Part A data suggests strategies using SUVmax or mean and considering SUV change in the hottest lesion may be optimal in TNBC and that a threshold change in excess 30-40% would be valid in the context of parameter repeatability.

8 Preliminary evaluation of FDG SUV response

8.1 Aims

Based on the repeatability results (Chapter 7) the FDG tracer was selected for progression to part B response evaluation. In all visualised tumour lesions SUVmax, SUV mean, SUVpeak and SULpeak parameters were evaluated and PET response considered with respect to the protocol defined mid-chemotherapy contrast-enhanced MRI RECIST primary endpoint as well as end of chemotherapy secondary endpoints (end of treatment MRI RECIST response and definitive RCB pathological response). This chapter presents the emerging response data for static scan acquisitions in the nine recruited patients who have completed scheduled Part B PET imaging response evaluations. Recruited participants who have not yet completed neoadjuvant chemotherapy are not considered in the analyses presented in this chapter.

8.2 Patent characteristics

Baseline patient and tumour characteristics for nine of the planned 20 patients that have completed neoadjuvant chemotherapy are summarised in Table 8.1. All primary tumours were confirmed ER and HER2 negative and were generally high grade NST. One patient had multicentric disease at the outset, the larger ER- HER2- primary defining her neoadjuvant treatment pathway and study participation. Only this lesion was considered for imaging and tissue evaluations purposes. All stated axillary nodal involvement was cytology confirmed. Scheduled neoadjuvant sequential chemotherapy comprised 4 cycles of docetaxel (T) 100mg/m^2 followed by 4 cycles of epirubicin 90mg/m^2 and cyclophosphamide 600mg/m^2 (EC). A single patient underwent surgery without transition to EC due to overt clinical progression despite stable mid-MRI response assessment.

Demographics		
Age (years)	Mean (s.d.)	44.5 (4.7)
	Median (min-max)	45.5 (35-49)
Race	Black	3 (33.3%)
	White	5 (55.5%)
	Asian	1(11%)
Body Mass Index	Mean (s.d.)	27.22 (5.5)
	Median (min-max)	27.2 (22-37.6)
Baseline Tumour characteristics		
Primary tumour	Mean (s.d)	34.8 (13.6)
(MRI longest dimension, mm)	Median (min-max)	29 (13-55)
Nodal involvement	Positive	5
	Negative	4
Number of lesions	1	4 (44.4%)
(breast and axilla)	2	5 (55.6%)
Histological type	NST	9 (100%)
Grade	2	1 (11%)
	3	8 (89%)
ER Score	0	7 (78%)
	1	0 (0%)
	2	2 (22%)
HER2 score	0-1	8 (89%)
	FISH -ve	1 (11%)
Chemotherapy (completed patients only, n=9)		
Administered T	3 cycles	1
	4 cycles	8
Administered EC	4 cycles	8

Table 8-1 Overall characteristics of recruited Part B patients.

Clinical characteristics of the nine recruited Part B patients who have completed all study investigations. Categorical variables are reported by frequencies and column percentages. s.d. = standard deviation. T = taxane, EC = epirubicin and cyclophosphamide neoadjuvant chemotherapy.

8.3 MRI response

Table 8-2 summarises MRI RECIST response following 3 cycles of docetaxel (mid-MRI) and end of sequential neoadjuvant chemotherapy (EOT-MRI). On the mid-MRI scan no patient exhibited progression by RECIST criteria however one participant underwent early surgery due to overt clinical progression following 3 cycles of docetaxel.

8.4 Pathological response

All patients have undergone definitive surgery. Breast conserving surgery (BCS) was performed in three patients and mastectomy (M) in the remainder. The 5 patients with initially node positive disease underwent axillary nodal clearance (ANC). Sentinel lymph node biopsy (SNB) was performed in the remainder. Contralateral prophylactic mastectomy was performed in one patient.

RCB scores are summarised in Table 8-2. Patient 3 presented with multicentric disease from synchronous ER- HER2- and ER+ HER2- primaries. Persisting disease at synchronous ipsilateral ER+ breast primary gave an overall RCB score of 2, RCB index 1.812. However the mastectomy specimen demonstrated no evidence of residual disease at the coil marked ER- HER2- breast cancer site (RCB0). As the imaging assessments for purposes of this study considered the TNBC primary only and ER+ RCB2 response predicts a substantially lower relapse risk than the same score arising from residual TNBC (76% vs. 54% 10 year RFS) (16) the RCB score pertaining to the TNBC cancer only is considered for comparative analysis .

The chi-squared test was applied to examine associations between mid-MRI response and later EOT-MRI and RCB response. No statistically significant association was seen between mid-MRI categories of ‘responder’ (complete and partial) or ‘non-responder’ (stable or progression) with EOT-MRI response category, p= 0.22 or with RCB categories of responder (RCB0-1) and non-responder (RCB2-3), p=0.13.

Patient	1	2	3*	4	5	6	7	8	9
MRI Imaging response									
Mid-MRI	CR	SD	PR	SD	PR	PR	SD	PR	PR
EOT-MRI	CR	-	PR	SD	CR	PD	PD	CR	PR
Pathological Response									
Tumour longest dimension (mm)	0	55	0	0	0	95	14	0	0
Involved Nodes (n)	0	0	0	2	0	8	1	0	0
RCB Index	0	2.672	0*	1.100	0	3.955	3.775	0	0
RCB Class	0	2	0*	1	0	3	3	0	0

Table 8-2 Summary of cross-sectional imaging and definitive Residual Cancer Burden response

*MRI response categorised according to RECIST 1.1 (31): CR = complete response; PR = partial response; SD = stable disease; PD = progressive disease. * denotes patient who had residual disease at the site of a second ER+ cancer at a separate location within the index breast (RCB2, index 1.812) but pathological response RCB0 to the separate TNBC lesion.*

8.5 PET scan acquisition parameters

The same scanner was used for image acquisition at each patient study visit. Where repeatability imaging was performed (Part A patients scanned with FDG), the second of the two pre-chemotherapy image acquisitions was taken as baseline for Part B response assessment. Chemotherapy commenced within median 4 (range 3-8) days of the baseline scan. Response PET imaging took place in the third week following cycle 1 in all study participants at median 18 (range 17-19) days.

All FDG scans were acquired after minimum 6 hours fast time. Follow up scan acquisition commenced within 4 minutes of the baseline acquisition at each of 90, 120 and 180 minute time points. No patient had serum glucose level greater than 6.2 mmol/l (normal range 4.4 to 6.1 mmol/L). The group mean serum glucose levels were 5.34 mmol/L (s.d. 0.43) and 5.1 mmol/L (s.d. 0.73) at the baseline and post cycle 1 visits respectively and did not statistically

differ between the scan visits (p=0.46). Mean injected activity was 183.1 (s.d. 5.9) MBq and 185.1 (s.d. 7.2) MBq for baseline and post cycle 1 scans respectively and did not statistically differ between visits (p=0.50). Response scan injected activity was within 10.4% (range -4% to 10.4%) of baseline injected activity for each participant. Comparison of group background SUL at baseline with the post cycle 1 response scan showed no significant difference at any time point (p=0.81, p=0.19, p=0.70 for 90, 120 and 180 minutes respectively). At each acquisition time point the group background SUL was within 20% of group baseline. However when considered on a per patient basis only the 90 minute acquisition met this PERCIST criterion (maximum 19% difference in background).

8.5.1 Lesion evaluability on baseline and follow-up PET scans

All scans were deemed suitable for analysis following visual inspection by an experienced nuclear medicine physician (MOD or SB) and scan analysis was performed by JG. The SUV parameters for baseline and post cycle 1 response scans are summarised in Table 8.3. 16 lesions were visualised on the baseline acquisitions, comprising all known breast primaries and involved axillary nodal lesions. A previously undiagnosed internal mammary chain (IMC) node was visualised in 1 patient and axillary nodal lesion in one further patient. All lesions were evaluable using SUVmax and SUVmean parameters at baseline 90, 120 and 180 minute acquisitions. Consistent with the Part A data, peak parameters could not be derived for all lesions within the 40% isocontour (Table 8-3) including breast tumours at 90 minutes (3 patients), 120 minutes (4 patients) and 180 minutes (3 patients) respectively. The 40% isocontour was selected based on existing data from other tumour sites which indicates the best size match between isocontour and pathological site is obtained using a 40% isocontour (65)(66, 67). It seems likely the evaluability of the Peak parameter is a function of the isocontour selected to define the metabolically active tumour size. In the current dataset peak lesion evaluability on baseline scans could be increased by applying a lower (e.g 20%) isocontour to define an expanded tumour VOI. Similarly reduced peak evaluability was observed where thr isocontour threshold to define tumour VOI was increased. PERCIST guidance proposes use of the 70% isocontour threshold to define tumour VOI however using this approach only a single participant had Peak evaluable tumour at baseline using this approach (patient 2, MRI tumour dimension 55mm)

On post cycle 1 PET imaging the 15 breast and axillary nodal lesions were visualised and evaluable using maximum and mean SUV parameters at all three acquisition time-points. It was not possible to visualise the IMC node. Fewer lesions were non evaluable using peak

parameters (Table 8-3). These lesions were not necessarily those evaluable on the baseline scan and consequently paired baseline and response peak parameters could be derived for a single lesion in only 4, 6 and 5 participants at 90, 120 and 180 minute acquisitions respectively.

FDG scan Acquisition		Baseline			Response		
		90	120	180	90	120	180
SUV max	Lesion n	16	16	16	15	15	15
	Mean (s.d.)	10.7 (5.9)	11 (6.3)	12.4 (7.3)	5.2 (3.3)	5.3 (3.7)	6.0 (4.7)
	Median	9.9	9.9	11.0	9.9	4.1	4.8
	Range	2.7-25.9	2.7-26.8	2.7-31.4	1-16	1.3-16.3	1.8-21.0
SUV mean	Lesion n	16	16	16	15	15	15
	Mean (s.d.)	6.4 (3.7)	7.0 (3.8)	7.7 (4.7)	3.0 (2.1)	3.2 (2.3)	3.5 (2.7)
	Median	6.0	6.7	2.9	6.7	2.4	2.9
	Range	1.6-16.8	2.6-17.8	1.6-20.4	1.0-10.0	1.0-10.1	1.0-12.0
SUV peak	Lesion n	10	9	9	7	8	7
	Mean (s.d.)	10.4(5.6)	10.5 (5.8)	12.6(6.3)	4.5 (3.7)	3.9 (4.1)	4.7 (5)
	Median	10.1	9.6	10.6	3.4	2.9	3.2
	Range	4.2-24.4	4.1-24.2	6.7-27.4	2.1-12.8	0.9-13.7	1.7-15.8
SUL peak	Lesion n	10	9	9	7	8	7
	Mean (s.d.)	7.1(3.9)	7.6(4.0)	8.5 (4.5)	3.1 (2.5)	2.7(2.7)	3.2 (3.3)
	Median	6.7	6.8	6.9	2.5	1.9	2.0
	Range	2.5-16.3	4.1-16.2	4.5-18.3	1.4-8.6	0.6-9.2	1.07-10.67

Table 8-3 Summary of standard uptake parameters for the FDG tracer

Results reported for group baseline and response scans at the 90, 120 and 180 minute static acquisition time points, all evaluable lesions for each parameter. Raw standard uptake values, mean, standard deviation (s.d.) and range shown for each parameter. SUVpeak and SULpeak could not be consistently derived on the baseline images at each of 90 minutes (5 lesions, 4 patients), 120 minutes (7 lesions, 6 patients) and 180 minutes (6 lesions, 5 patients).

8.6 SUV response

Table 8.4 reports the % change in SUVmax and SUVmean parameters at the post cycle 1 time point. In 3 of the 9 participants axillary lesions exhibited the highest baseline uptake, meeting hottest lesion criteria for response monitoring. The observed % decrease in target lesion (breast or axilla) SUVmax was in the range -39.9 to -76.7%, -37.2 to -82.4% and -33 to -84.5% at each of the 90, 120 and 180 minute scan acquisitions respectively irrespective of later radiological or pathological response. Similarly, % change in SUVmean was in the range -42.8 to -79.2%, -39.4 to -81.8% and -41.3 to -86.9% at 90, 120 and 180 minute acquisitions respectively. Thus all target lesions exceeded 15-30% SUV change and would be classified as metabolic responders using PERCIST and EORTC criteria.

Patient	Lesion	% change SUVmax			% change SUVmean			Mid-MRI response	RCB
		90	120	180	90	120	180		
1	**Breast**	-76.7	-82.4	-84.5	-79.2	-81.8	-86.9	CR	0
2	**Breast**	-39.9	-39.3	-33.0	-42.8	-43.1	-41.3	SD	2
3	**Breast**	-52.5	-51.1	-55.6	-55.4	-54.9	-52.7	PR	0
4	Breast	-50.9	-43.5	-50.7	-54.5	-46.1	-52.7	SD	1
	Axilla	-57.0	-59.5	-57.0	-57.2	-58.8	-56.2		
5	**Breast**	-46.9	-51.7	-52.8	-48.6	-52.2	-53.8	PR	0
	Axilla	-23.9	-36.3	-19.6	-27.2	-37.7	-16.2		
6	**Breast**	-44.1	-39.4	-45.2	-46.3	-39.4	-44.2	PR	3
	Axilla	-57.4	-58.4	-55.6	-60.8	-59.5	-56.7		
7	Breast	-50.7	-50.0	-48.9	-56.3	-57.3	-56.5	SD	3
	Axilla	-44.5	-37.2	-40.8	-47.3	-40.2	-45.3		
	IMC	-	-	-	-	-	-		
8	Breast	-52.7	-60.8	-58.4	-52.9	-60.7	-58.9	PR	0
	Axilla	-54.9	-47.9	-59.4	-55.0	-48.2	-60.9		
9	**Breast**	-72.4	-76.3	-75.8	-73.9	-76.0	-77.3	PR	0
	Axilla	-46.9	-50.7	-35.2	-49.7	- 37.7	-39.6		

Table 8-4 SUVmax and SUVMean response following 1 cycle of chemotherapy

Per lesion % change in SUVmax and SUVmean parameters shown for the 90, 120 and 180 minute scan acquisitions. Where patients presented with more than one visible lesion, the hottest lesion target is denoted in bold. The IMC node (patient 7, non target lesion) was visualised on the baseline acquisitions only, thus % change in max and mean parameters could not be calculated for this lesion.

Non-evaluability of peak parameters within the 40% isocontour on both baseline and response scans meant % change could only be calculated in 4, 5 and 6 lesions at the 90, 120 and 180 minute acquisitions respectively. This precludes meaningful interpretation of response using SUV and SULpeak in the current cohort and if confirmed at full study accrual indicates that neither peak parameter is suitable for application as a response biomarker in TNBC.

8.6.1 Primary endpoint comparison of % change in SUV with mid MRI response

Mid-MRI RECIST response was categorised as 'responder' (complete and partial; n=6 patients with total 10 lesions) or 'non-responder' (stable or progression; n=3 with total 5 tumour lesions) for comparison with % SUV change. The distribution in % SUV change did not statistically differ between categories of mid-MRI response when considering all lesions (p=0.56, 0.21 and 0.38 for SUVmax and 0.64, 0.46 and 0.64 for SUVmean at 90, 120 and 180 minutes scan acquisitions respectively) (Figure 8-1i). Similar findings were present when considering hottest lesion targets only (Figure 8-1ii) (p=0.27, 0.29 and 0.11 for SUVmax and 0.26, 0.32 and 0.18 for SUVmean at 90, 120 and 180 minutes scan acquisitions respectively).

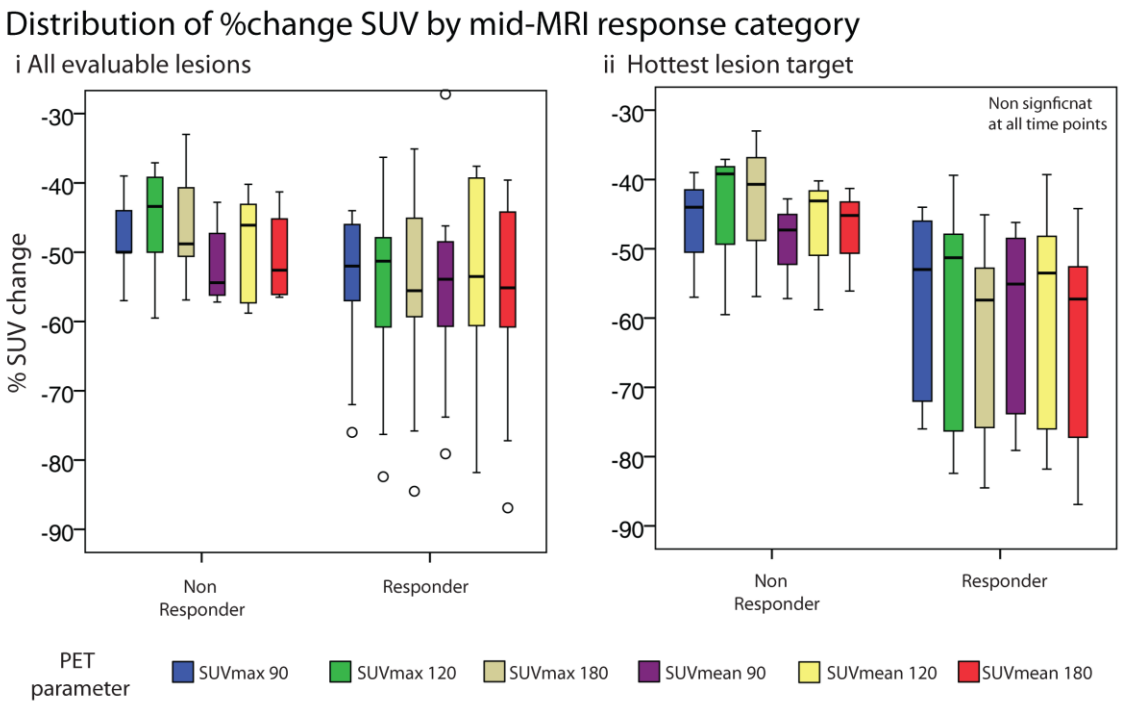


Figure 8-1 Distribution of % SUV change according to mid-MRI response category

i. Mid-MRI RECIST response was categorised as 'responder' (complete and partial, n=6 patients with total 10 lesions) or 'non responder' (stable or progression, n=3 with total 5 tumour lesions) for comparison with % SUV change. SUVmax and SUVmean parameters were evaluated for tumour lesions visualised on the 90, 120 and 180 minute static acquisitions. Box plots show comparison of distribution for % SUVchange in categories of mid-MRI Responder and Non responder (i) all lesions, n=15 lesions and (ii) hottest lesion target only, n=9 lesions. No significant difference in distribution of % SUV change was present for any parameters or acquisition time point. Outlier lesions (i) were from patients 1 and 5, both of whom later achieved complete radiological response on the EOT-MRI scan and RCB0 definitive pathological response

8.6.2 Secondary endpoint comparison of % change with EOT-MRI and RCB response

EOT-MRI RECIST response was categorised as 'responder' (complete and partial; n=6 patients with 8 lesions) or 'non-responder' (stable or progression; n=2 patients with 4 lesions) for comparison with % SUV change (Figure 8-2 Ai). The distribution of % SUV change did not significantly differ not between categories of EOT-MRI response when considering all lesions (p=0.61, 0.48 and 0.24 for SUVmax and 0.46, 0.73 and 0.64 for SUVmean at 90, 120 and 180 minutes scan acquisitions respectively).

Considering the hottest lesion target only, inspection of box and whisker plots suggests separation between the two categories of EOT-MRI response. Patient 3 underwent early

surgery due to clinical progression after only 3 cycles of chemotherapy but had SUV values within range of the non-responding group for all parameters and acquisition time points (Figure 8-2 Aii). Considering hottest lesion target only significant differences in distribution of % SUV change between categories of EOT-MRI were most marked for the 120 and 180 scan acquisition (p=0.048, 0.034, 0.019 for SUVmax and p=0.063, 0.037 and 0.043 for SUVmean at 90, 120 and 180 minutes respectively) in the current cohort. Full accrual will be required to definitively comment on the statistical significance of differences between the two categories.

RCB score was categorised as 'responder' (RCB 0 or 1; n=6 patients with 10 lesions) or 'non-responder' (RCB 2 or 3; n=3 patients with 5 lesions) for comparison with % SUV change. Considering all lesions, no significant differences in the distribution of % SUV change between responding and non-responding categories was present for SUVmax and SUVmean at any timepoint (Figure 8-2Bi) (p=0.37, 0.14 and 0.18 for SUVmax and 0.48, 0.21 and 0.38 for SUVmean at 90, 120 and 180 minutes scan acquisitions respectively). Considering hottest lesion targets only, inspection of box and whisker plots suggests separation between the two groups that is more marked at 120 and 180 time points for both SUV parameters. Significant differences in distribution of % SUV change is present between categories of RCB response, $p<0.05$ at all scan acquisition times, (p=0.048, 0.012 and 0.019 for SUVmax and p= 0.022, 0.012 and 0.014 SUVmean at each of 90, 120 and 180 minutes scan acquisitions) (Figure 8-2Bii).

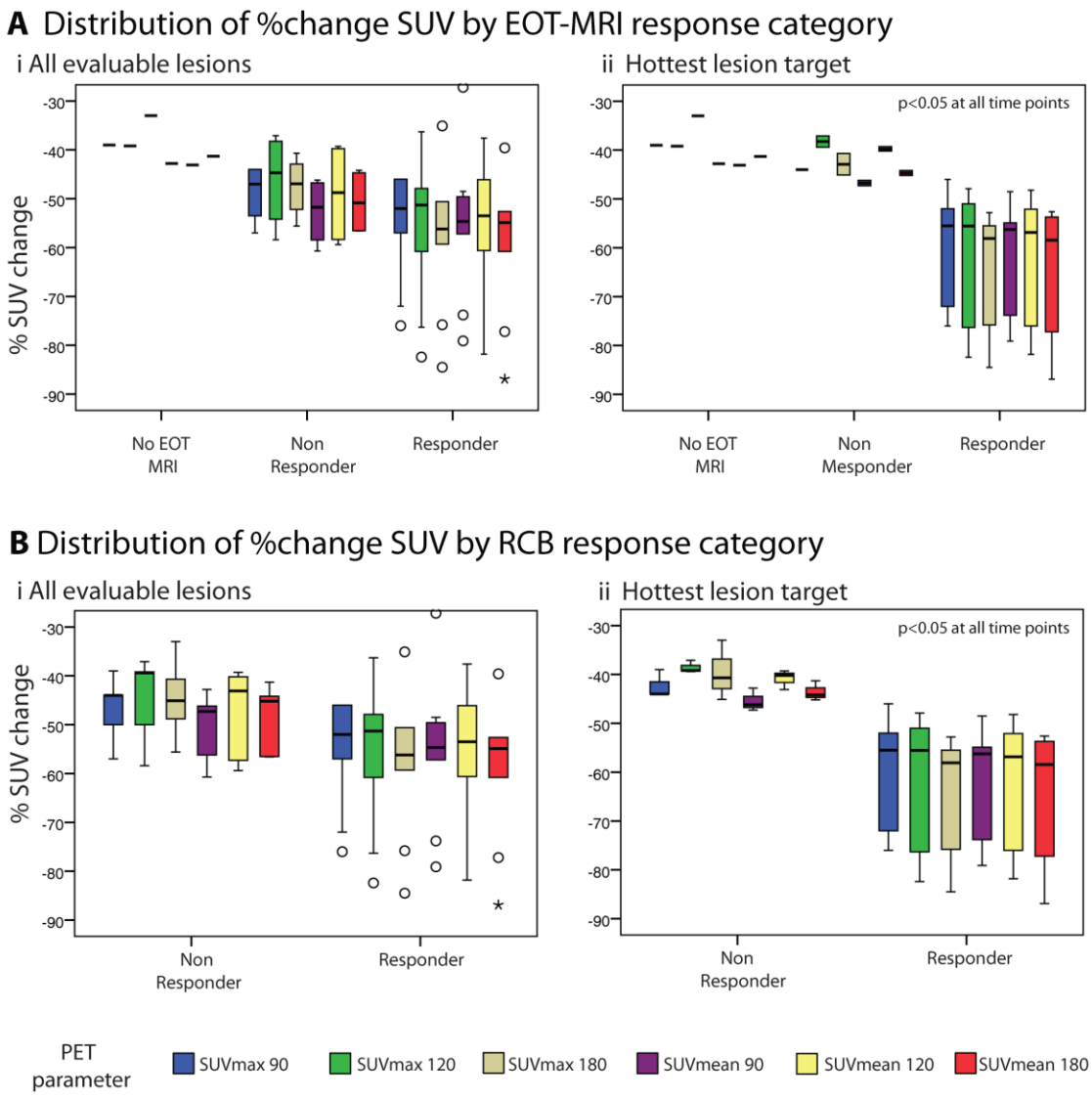


Figure 8-2 Distribution of % SUV change according to EOT-MRI and RCB response categories

A. Box plots showing comparison of distribution for % change in SUVmax and SUVmean in categories of EOT-MRI response (i) all lesions, n=15 and (ii) hottest lesion target only (n=9). Significant difference in distribution of % SUV change is present when considering hottest lesion target for SUVmax at each of 90, 120 and 180 minute FDG scan acquisitions and for SUVmean at 120 and 180 minutes. Patients with outlier lesions (i) subsequently achieved pathological complete response, RCB0 at definitive surgery. % SUV change also shown for the single patient underwent early surgery following the mid MRI due to overt clinical progression and therefore did not receive EOT-MRI.

B. Box plots showing comparison of distribution for % change in SUVmax and mean in categories of RCB response (i) all lesions (n=15) and (ii) hottest lesion target only (n=9). Considering hottest lesions only, significant differences in distribution of % SUV change are present at 90, 120 and 180 minute FDG scan acquisitions (p<0.05 at each time point).

8.7 Pre neoadjuvant SUV assessment

The distributions of baseline SUVmax and mean parameters were compared with categories of mid-MRI, EOT-MRI and RCB response. No significant differences in distribution were present at any FDG scan acquisition time-point ($p>0.1$ for all comparisons).

8.8 Research Tissue Evaluations

Biopsy derived biomarkers such as ki-67 may have prognostic and predictive potential (36) and potentially therapy induced changes in the tumour correlate with subsequent survival outcome (37) and may add to RCB (38). However there is little data considering relationship of these metabolic changes with FDG PET derived SUV indices. To address these questions research biopsies were required with the study protocol to enable exploratory correlation of biopsy derived markers of proliferation, apoptosis and glucose transport with PET imaging (baseline SUV and post cycle 1 SUV change) and later pathological response at definitive surgery.

Research breast core biopsy specimens from the 11 patients (part A and B cohort) who had completed neoadjuvant chemotherapy and definitive surgery by August 2015 were examined to confirm presence of tumour (Professor Sarah Pinder, SP). In tumour containing samples Ki-67, Geminin, Activated Caspase 3, Mini-chromosome maintenance protein 2 (MCM2) and Glucose transporter 1 (GLUT-1) staining assessment was performed and scored as a percentage of positive cells within the tumour cell population (PC and SP) using standard methods (88, 168-170).

The raw data for both FDG and FLT imaged participants (Table 8-5) were considered for RCB comparisons. Baseline and post cycle 1 biopsies contained evaluable tumour in 7 of the 11 patients. Comparison of baseline and post cycle 1 scores revealed statistically significant differences for GLUT-1 only (Wilcoxon Signed Rank Test). As expected the on-treatment samples were less likely to be evaluable in those achieving RCB 0 or 1 at definitive surgery. The three individuals with tumour containing samples at each of baseline, post cycle 1 and end of chemotherapy time points achieved RCB 2 or 3 definitive histology response. For those with poor pathological response (RCB 2 or 3) inspection of the data provided in Table 8.5 provides little evidence of tissue biomarker change compared to baseline between RCB responding and non-responding subsets. Application of the Mann Whitney U test to compare the distribution of Ki-67, cleaved caspase 3, geminin, GLUT-1 and MCM2 between RCB responding and non-responding subsets identified no statistically significant differences for any parameter at baseline ($p=0.56$, $p=0.11$, $p=0.56$, $p=0.73$ and $p=0.2$ respectively) or the post cycle 1 time point

(p=0.86, p=0.56, p=0.40, p=1 and p=0.5 respectively). Similarly in the 7 patients with evaluable tissue samples at baseline and post cycle 1 time points no statistically significant difference in distribution of % change in Ki-67, cleaved caspase 3, geminin, GLUT-1 and MCM2 was present between responding and non-responding patients (p=0.63, p=0.1, p=0.63, p=0.63 and p=1 respectively). No patient achieving RCB 0 or 1 had residual tumour in the breast precluding tissue evaluation at this time-point. The validity of statistical comparison is with the currently small sample size is highly constrained and the full dataset will be required for meaningful biopsy derived biomarker comparisons.

In this dataset research tissue evaluations did not exhibit normal or log normal distributions precluding linear regression analysis for comparison with log transformed SUVs. Correlations between the pre-chemotherapy SUV parameters (FDG imaged breast lesions only) and baseline research tissue evaluations were explored using the Spearman correlation coefficient. At all scan acquisition time points statistically significant correlation was present at the p=0.05 level between SUV (max and mean) and the Ki-67 and MCM2 parameters only (Table 8-6). However as previously noted further data is required for objective comparison.

	Patient	FDG Group							FLT Group				Median	*p (comparison with baseline)*
		1	**2**	**3**	**4**	**5**	**6**	**7**	**1**	**2**	**3**	**4**		
Ki-67	Baseline	75.65	88.53	37.12	23.20	-	79.59	85.00	87.5	85.24	65.00	-	79.59	
	1 cycle	-	94.33	20.38	-	-	65.58	80.00	75.00	85.10	60.00	-	75	*0.90*
	EOT	-	84.23	-	-	-	75.00	60.00	-	-	-	35.00	67.5	*0.11*
Geminin	Baseline	16.31	32.92	10.90	13.61	-	21.29	28.75	37.71	0.03	24.34	-	21.29	
	1 cycle	-	36.64	5.26	-	-	11.92	31.73	34.26	0.01	28.83	-	21.83	*0.18*
	EOT	-	26.00	-	-	-	4.65	18.78	-	-	-	6.79	12.78	*0.11*
Cleaved Caspase 3	Baseline	0.65	0.315	0.27	1.18	-	1.18	1.3	1.19	0.62	0.88	-	0.88	
	1 cycle	-	0.27	0.38	-	-	1.15	1.22	1.25	0.43	1.09	-	1.09	*0.87*
	EOT	-	0.285	-	-	-	0.34	0.65	-	-	-	0.00	0.31	*0.11*
MCM2	Baseline	62	62.38	12.9	5	-	47.44	73.37	58.03	52.03	36.85		52.03	
	1 cycle	-	43.48	11.5	-	-	24.88	85.34	45.00	36.85	-	-	40.17	*0.08*
	EOT	-	49.81	-	-	-	25.07	60.02	-	-	-	6.21	37.44	*0.11*
GLUT-1	Baseline	69.02	37.63	4.38	4.77	-	15.80	71.63	42.69	31.69	27.09	-	31.69	
	1 cycle	-	2.34	3.24	-	-	2.63	21.19	14.80	0.10	0.75	-	2.63	*.02*
	EOT	-	33.19	-	-	-	41.25	52.32	-	-	-	1.15	33.22	*0.11*
RCB Class		0	2	0*	1	0	3	3	0	0	2	2		

Table 8-5 Tissue analyses for evaluable samples (Part A and B cohorts)

Ki-67, Geminin, Cleaved Caspase 3 , GLUT-1 and MCM2 scored as a percentage of positive cells within the tumour cell population. Results from evaluable baseline, post cycle 1 chemotherapy and end of treatment (EOT) samples in patients who completed neoadjuvant and definitive surgery (combined patient from Parts A and B. – denotes non evaluable sample. Group comparison of baseline and on treatment tissue scores revealed statistically significant differences for GLUT-1 at the post cycle 1 time point only (Wilcoxon Signed Rank Test)

		90 minute		120 minute		180 minutes	
		SUVmax	**SUVmean**	**SUVmax**	**SUVmean**	**SUVmax**	**SUVmean**
Ki-67	Correlation coefficient	0.83	0.083	0.77	0.83	0.83	0.83
	p	0.04	0.04	0.07	0.04	0.04	0.04
Geminin	Correlation coefficient	0.77	0.77	0.71	0.77	0.77	0.77
	p	0.07	0.07	0.11	0.07	0.77	0.77
Cleaved Caspase 3	Correlation coefficient	-0.15	-0.15	-0.23	-0.15	-0.15	-0.15
	p	0.78	0.78	0.66	0.78	0.78	0.78
GLUT-1	Correlation coefficient	0.12	0.12	0.36	0.12	0.12	0.12
	p	0.82	0.82	0.94	0.82	0.82	0.82
MCM2	Correlation coefficient	0.83	0.83	0.66	0.83	0.83	0.83
	p	0.04	0.04	0.16	0.04	0.04	0.04

Table 8-6 Correlation of baseline breast SUV with tissue evaluations (FDG cohort)

Results from FDG imaged patients with evaluable baseline tissue samples (n=6). Ki-67, Geminin, Cleaved Caspase 3 and GLUT-1 scored as a percentage of positive cells within the tumour cell population. Significant correlation is present for Ki-67 and MCM2 with baseline SUVmax and SUVmean parameters only (Spearman's Rank correlation coefficient, 2 sided)

8.9 Discussion

Part B seeks to relate changes seen on post cycle 1 PET to later conventional imaging and pathological response. The tempo of recruitment has been much slower than hoped due to the tracer production and scan delivery challenges that slowed completion of part A and the comparative rarity of TNBC. Despite this, strengths of the current PET data set include the clean TNBC phenotype and representation of the full spectrum of pathological response following neoadjuvant therapy. EANM control/quality assurance procedures concerning consistency in patient preparation, scan procedure and image reconstruction for FDG image interpretation have been tightly adhered to (69) and consistency of scan scheduling in relation to therapy and acquisition time following tracer injection achieved. PERCIST criterion (62) relating to group background activity at the 90 minute scan acquisition were met. Nevertheless a significant proportion of patient lesions were not evaluable using SUL and SUV peak measures. In this dataset tumour VOI was defined using a 40% isocontour and increasing the isocontour threshold above 40%detrimentally impacted on lesion evaluability using Peak parameters. PERCIST guidance (62) suggests a 70% threshold, however using this method only

a single patient had peak evaluable tumour at baseline (patient 2, clinical dimension 55mm). Although the current sample size is small, it has clinical parameters representative of the TNBC population for whom neoadjuvant therapy is commonly recommended (171). An alternative study design that excluded those with primary breast tumours at the lower end of the T2 (2-5cm) spectrum may have permitted higher rates of lesion evaluation using peak measures but this approach risks inherently constraining the future clinical population in whom the PET response data might be applied as well as further slowing study accrual. In contrast these data confirm SUVmax and SUVmean parameters are evaluable in all biopsy or cytology proven lesions and in this regard better meet the needs of the neoadjuvant TNBC population.

Study recruitment continues but within the current dataset there is no indication that % SUV change differs between categories of the primary end-point mid-MRI response. This mirrors the lack of a clear relationship observed between mid MRI response and later EOT-MRI or pathological outcome. However the emerging data suggests magnitude of % change in SUV (max and mean) may differ between categories of later EOT-MRI and RCB response. Furthermore the % change in hottest lesion target, either breast or axilla, appears to provide the best relationship with response at completion of neoadjuvant chemotherapy. In the context of a clinical requirement for early identification of poorly responding sub-populations who may derive most benefit from participation in clinical trials of novel agents, this preliminary data provides a strong rationale for completing recruitment. Considering hottest lesion targets only, significant differences in distribution of % SUV change between categories of RCB response are present, but the full accrual will be required to properly comment on superiority of delayed image acquisition or the optimal SUV parameter. In addition the magnitude of SUV change present in those with later poor pathological response raises the possibility that the relatively tighter repeatability present at the 90 minute time point (Chapter 7) may be an important factor in defining the most predictive acquisition time and SUV parameter.

FDG PET performance is recognised to be dependent on breast cancer subtype (172). Despite the mix in subsequent pathological response, all target lesions in the current dataset exhibited greater than 30% drop in SUV at the post cycle 1 time point thus meeting existing criteria to define metabolic response (62). Groheux *et al.* suggest as high as 42% threshold drop (SUVmax) is required to differentiate between those who will achieve pCR from those with residual TNBC at the post cycle 2 time point (56). More data will be required prior to receiver operator curve assessment or consideration of subcategories of RCB in the TNPET dataset. However it seems likely discriminatory threshold for later RCB0/1 level of response will similarly exceed the 30% level at the earlier post cycle 1 time point further supporting

requirement for TNBC specific response criteria. The current promising early data suggests FDG PET response assessed using change in SUVmax or SUVmean will be predictive at the clinically desirable post cycle 1 timepoint however completions of the planned accrual will increase confidence in this finding and may inform future validation studies.

The protocol design included tissue sampling at time points contemporaneous with PET imaging and at definitive surgery to permit comparison of tracer performance with biopsy derived indices of cell proliferation, apoptosis and glucose metabolism. Consistent with prior data in other high proliferation malignancies (173, 174) and in unselected breast cancer (44, 46) the cell proliferation-associated nuclear antigen, Ki-67, demonstrates significant correlation with SUV. Raw Ki-67 values have been presented (Table 8-5), but it should be noted that all samples in the current cohort exceeded a threshold staining level of 10-20% used to define high proliferation (88). Such high Ki-67 across the study cohort is consistent with published TNBC datasets (175) where average Ki-67 is in the range 35-48% (176-178) and only 11% samples have values below 20% (178). It is therefore unlikely Ki-67 in our dataset is an artefact of technical limitations associated with biopsy rather than whole tumour specimen assessment (87) or the low overall sample size.

In unselected breast cancer high Ki-67 at completion of neoadjuvant chemotherapy strongly predicts poorer outcome in patients not achieving a pathological complete response, particularly in ER+ subsets (37, 175). However the prognostic ability of Ki-67 alone is confounded by the underlying molecular subtype (179) and intuitively tumour cells that retain high proliferation at completion of treatment are also likely to be those most resistant to chemotherapeutic agents. Consistent with this, four of the five individuals achieving RCB 2 or 3 exhibited persisting high Ki-67 at completion of neoadjuvant therapy, two of whom have already developed metastatic relapse (FLT patient 3 and FDG patient 6), the latter within a few months of surgery. In contrast, lack of residual primary breast cancer at the end of treatment meant Ki-67 could not be reported for those achieving RCB 0 and 1 (FDG patient 4, single node residual disease). Although the current sample size is small, pCR rates in the breast are recognised to be higher in TNBC compared to other breast cancer phenotypes (5, 17), consequently an earlier on treatment biopsy may be required for Ki-67 change to be a meaningful biomarker. However the current data provides no suggestion that distribution of baseline Ki-67 nor change after 1 cycle significantly differs according to later RCB response; indeed two of the seven patients with evaluable tumour samples had very high Ki-67 at the post cycle 1 time point but subsequent RCB0 response. This may simply be an artefact of the current small sample size, technical issues associated with biopsy sampling (87), inter-lesion as well as intra-lesion heterogeneity of Ki-67 expression and differential impact of therapy on

breast and involved nodal disease. However if full study accrual confirms relationships between SUV change in hottest lesion target (breast or axilla site) and later RCB response but not Ki-67 it is plausible that some of the challenges associated with early tissue acquisition for response assessment may be mitigated by ability to 'sample' the whole tumour burden using PET imaging.

Other markers of proliferation have been identified as participants in the process of DNA replication and may have prognostic value. Mini-chromosome maintenance protein 2 (MCM2), is present in all phases of the cell cycle and detects cells capable of initiating DNA replication. It has been reported to be a more sensitive marker of proliferation than Ki-67 or geminin and in an unselected breast cancer cohort baseline MCM2 provided a more robust and sensitive prognostic marker for breast cancer specific survival (170). Geminin, expressed from S to M phase, identifies the sub-fraction that have entered S phase, but not exited mitosis thereby providing a more specific indicator of the fraction of cells that are replicating. Increased expression has been identified as an independent indicator of adverse prognosis in unselected breast cancer, predicting both poor overall survival and the development of distant metastases (169). Published data suggests MCM2 expression exceeding 12% defines high proliferation (170). In the current dataset this threshold was exceeded in 10 of 11 baseline samples (Table 8-5) indicating high levels of proliferation within the cohort. Published thresholds for defining high geminin are more dispersed, varying from 2.3 to 30% (170, 180) thus in this preliminary TNPET dataset between 2 and 7 of the 8 evaluable baseline samples exhibited high expression. Correlation of baseline geminin scores with FDG SUV parameters approached significance at the earlier (90 minute) scan acquisition (spearman's correlation coefficient 0.77, p=0.07) and significant correlation was present for MCM2 scores (spearman's correlation coefficient 0.83. p=0.04). However the current data provides no indication of statistically significant difference in the distribution of geminin or MCM2 between RCB responding and non-responding subsets at either baseline or the post cycle 1 time point nor significant difference in distribution of % change in these tissue biomarkers at 1 cycle. Apoptosis mediated by either the intrinsic or extrinsic pathways results in cleavage of caspase-3 and assessment of the activated caspase-3 antibody staining is a validated marker of apoptosis in breast cancer (168). However the current TNPET data provides no indication that baseline, post cycle 1 or on treatment change in activated caspase-3 antibody differs between categories of RCB good and poor response.

Published data reports association between FDG uptake in breast and GLUT-1 expression (42), with greatest expression levels in TNBC (177) or IHC defined basal phenotype (54). We are not aware of published clinical data reporting impact of cytotoxic therapy on GLUT-1 receptor expression for any tumour site but preclinical data suggests successful anticancer therapy

induces reduction in glycolysis that may be explained by downregulation of GLUT-1 expression (166) or loss of membrane GLUT-1 localisation (165). In the current cohort, group comparison of baseline and post cycle 1 scores identifies statistically significant difference in GLUT-1 expression. However high heterogeneity in intensity and extent of staining across whole tumour specimens is recognised (42) and may contribute to lack of correlation observed between pre-chemotherapy SUV parameters (FDG imaged breast lesions only) and GLUT-1 expression as well as the lack of significant difference in distribution of the staining level at baseline, post cycle 1 time point or the % change in GLUT-1 expression in the current data. Accrual of the full cohort will be required to more definitively comment on this.

In summary the emerging Part B response data suggests FDG PET response assessed using change in SUVmax or SUVmean derived from the hottest breast or axillary lesion does not differ between categories of mid-MRI response but will be predictive of later RCB response at the clinically desirable post cycle 1 time point. As early differentiation between categories of later good (RCB0/1) and poor (RCB2 and 3) response is of greater clinical value than prediction of the mid MRI response this finding is potentially very important. However completion of the planned accrual will be required to increase confidence in this finding and inform future validation studies before PET imaging response adaptive approaches can be progressed in TNBC.

9 Preliminary Evaluation of FDG Dynamic Imaging in TNBC

9.1 Objectives

Clinical PET-CT imaging is usually performed using static imaging and SUV evaluation over tumour sites. However full kinetic approaches, as they use the entire (measured) arterial input and tumour time-activity curves (TAC) in combination with a tracer kinetic model and plasma glucose levels to delineate both the temporal and spatial pattern of tracer uptake, are considered the most accurate quantitative measures and potentially provide greater information about in vivo tumour biology.

This chapter reports evaluation of the dynamic scan acquisitions using the FDG tracer to assess early docetaxel response in TNBC. The dynamic parameters of overall influx rate constant (K_i) in min^{-1} and the metabolic rate of glucose (MRGlu) in $mmol \cdot l^{-1} \cdot min^{-1}$, which equals K_i times blood glucose have previously been reported as potentially superior to SUV for *in vivo* interrogation of malignant tumours (181) and were therefore selected for evaluation using Patlak and FDG two tissue compartment modelling (FDG TCM). Repeatability estimates and preliminary response evaluation data for the nine FDG imaged patients who have completed research imaging and undergone definitive surgery are presented.

9.2 FDG Dynamic Scan Evaluation Methods

The anonymised DICOM PET and CT image files for FDG imaged participants were imported from the HERMES workstation for quantitative dynamic interpretation using PMod3.4 (PMOD Technologies Ltd, Zurich) according to the method described (6.5.3). Scan acquisitions in each individual were performed using the same scanner at all time points, but the GSTT PET Centre refurbishment required a change of scanner after the first three FDG imaged participants. PMOD assumes that all data loaded for kinetic modelling is already decay corrected unless in-programme correction is performed. Initial evaluation of the time activity curves (TAC) generated within PMOD indicated need for decay correction of scans acquired using the GE Discovery VCT 64 slice PET-CT. To verify the PMOD 3.4 decay correction process FDG phantom DICOM datasets acquired using the VCT PET-CT with and without on-scanner decay corrected image reconstruction were loaded for kinetic analysis (four 30 minute dynamic acquisitions performed using a cylindrical phantom over 6 hours, provided courtesy of PET physicist Lucy Pike). Phantom time activity curves (TACs) derived from the dataset decay corrected within PMOD ('operations' tool and inputting the 18F half-life, 109.8 minutes) exactly matched the on scanner decay corrected dataset (Figure 9.1) confirming suitability of this method for decay

correction of the clinical image data prior to analysis. Scans acquired post refurbishment using the GE Discovery 710 were decay corrected on the scanner and did not require within PMOD decay correction.

Using decay corrected imaging data, tumour and blood input volumes of interest (VOI) were defined within PMOD 3.4 on each dynamic FDG acquisition and Ki and MRGLu values for each breast tumour using the image derived arterial blood input using Patlak and FDG 2TCM models as described (Section 6.5.3). Measured blood activity derived from a single venous sample acquired at 60 minutes post tracer injection permitted rescaling of the arterial input by multiplying the kinetic parameter by ratio of the manual blood sample to the time matched descending aorta region of interest in 7 of the 8 imaged participants.

To facilitate assessment of SUV change over time, dynamic PET acquisitions obtained using the GE Discovery 710 64 slice PET-CT scanner (n=5) were reconstructed to provide four sequential static datasets at 40-50, 50-60, 60-70 and 70-80 minutes following tracer injection. SUV evaluation was performed using HERMES Hybrid Viewer Software version 1.4C Hermes, as previously (Section 6.5.3).

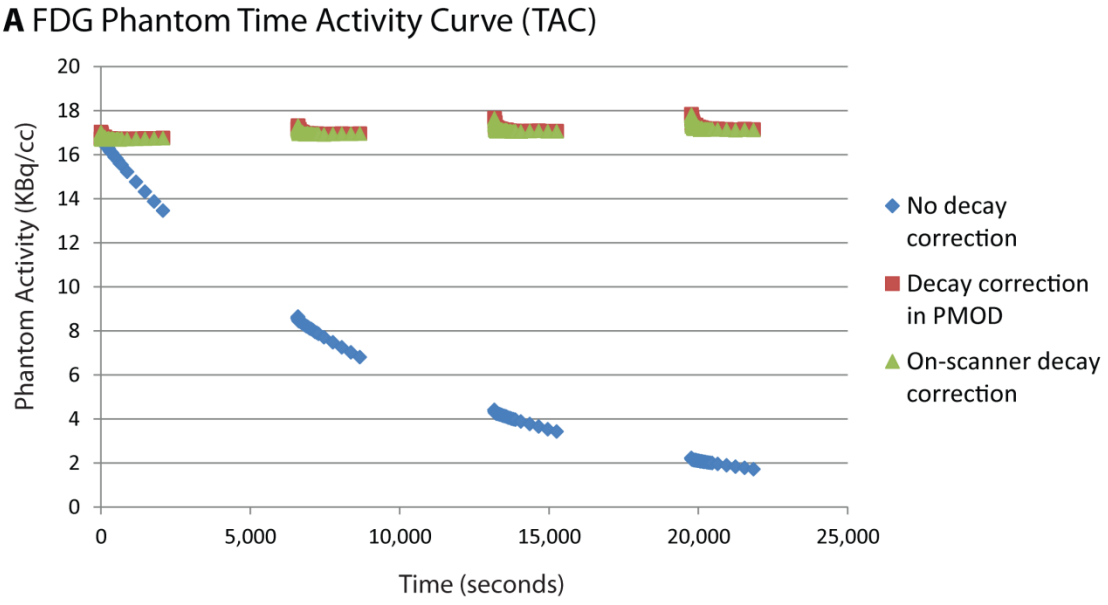


Figure 9-1 Validation of PMOD decay correction using FDG Phantom dataset

Verification of PMOD decay correction using a phantom dataset reconstructed with and without on scanner decay correction. Match between TACS generated using on-scanner and within P-MOD methods was confirmed

9.3 Dynamic results

Breast primaries could be visualised on dynamic acquisitions in all patients but as expected the 15 cm field of view precluded evaluation of involved nodal lesions. Both FDG two tissue compartment (2TCM) and Patlak models were applied for dynamic scan interpretation. Table 9.1 summarises group results for Ki and MRGlu parameters assessed at baseline (2 scans for Part A, 1 scan Part B) and at post cycle 1. The raw values (9 participants; 23 scans) and blood rescaled values (8 of 9 total participants, 21 scans) are presented. A wide dispersal of individual scan scaling factors is present (2TCM mean 1.12, range 0.72-2.14; Patlak mean 1.23, range 0.73-2.08), consistent with recognised limitations associated with use of image derived blood input including impact of partial volume effect, spillover to hot regions and patient movement. Consequently it was not appropriate to apply a generalised scaling factor to the single patient without a 60 minute venous activity sample.

		Ki			MRGlu		
		1st Baseline*	**2nd Baseline**	**Response**	**1st Baseline***	**2nd Baseline**	**Response**
2TCM (scaled)	**n**	**5**	**8**	**8**	**5**	**8**	**8**
	Mean	0.018	0.020	0.007	9.94	12.01	3.52
	s.d	0.019	0.016	0.009	10.09	11.04	4.39
	Median	0.011	0.016	0.003	7.07	8.83	1.90
	Min-max	0.004-0.052	0.003-0.056	0.0001-0.028	1.83-27.55	1.96-37.33	0.04-13.72
2TCM (raw)	***n***	***5***	***9***	***9***	***5***	***9***	***9***
	Mean	*0.019*	*0.018*	*0.008*	*10.05*	*11.54*	*4.00*
	Median	*0.016*	*0.013*	*0.004*	*8.01*	*7.31*	*2.63*
	s.d	*0.017*	*0.015*	*0.012*	*8.87*	*15.37*	*6.10*
	Min-max	*0.004-0.047*	*0.003-0.058*	*0.00004-0.003*	*1.71-25.01*	*1.71-38.78*	*0.02-14.73*
Patlak (scaled)	**n**	**5**	**8**	**8**	**5**	**8**	**8**
	Mean	0.018	0.015	0.006	9.83	10.93	3.23
	Median	0.014	0.011	0.003	6.96	6.42	1.51
	s.d	0.017	0.018	0.009	9.00	6.42	4.84
	Min-max	0.04-0.048	--0.011-0.051	-0.001-0.028	1.75-25.21	1.43-34.31	-0.03-14.04
Patlak (raw)	***n***	***5***	***9***	***9***	***5***	***9***	***9***
	Mean	*0.018*	*0.014*	*0.006*	*9.86*	*8.38*	*3.14*
	Median	*0.014*	*0.010*	*0.003*	*8.43*	*5.57*	*2.08*
	s.d	*0.015*	*0.021*	*0.013*	*7.82*	*14.28*	*6.28*
	Min-Max	*0.004-0.043*	*-0.011-0.053*	*-0.004-0.03*	*1.81-22.94*	*0.02-35.56*	*-0.02-15.02*

Table 9-1 Ki and MRGlu

Group Ki and MRGlu derived using FDG 2-tissue compartment model (2TCM) and Patlak models. In eight of the nine participants a venous activity sample at 60 minutes was available permitting rescaling of Ki and MRGluc parameters. Two baseline scan acquisitions were

performed in Part A participants. Mean, standard deviation (s.d.) median and range for raw and scaled kinetic parameters are presented. ** First baseline in Part A imaged patients only*

9.3.1 Repeatability of Dynamic Parameters

Ki and MRGlu parameters were confirmed to follow log normal distribution in the current dataset. For repeatability end point analysis (Part A participants only, n=5) Bland-Altman plots were constructed using the log transformed values derived using FDG two tissue compartment and Patlak models with and without rescaling using the venous activity sample (Figure 9-2 A and B). Overall Ki and MRGlu values derived using the rescaled 2TCM model exhibited tighter repeatability than values derived from either model with solely the image derived blood input or Patlak with rescaling but both parameters present breast lesions with differences of greater than 15% (0.14 on the log scale). Using the 2TCM, rescaled Ki values did not differ by more than 22% on the two scans (±0.2 on the logarithmic scale) and greater variability was present using MRGlu values (Figure 9.2 Ai and ii respectively). Rescaled Ki and MRGlu values derived using the Patlak model performed less well, differing by more than 50% (±0.41 on the log scale) even with rescaling.

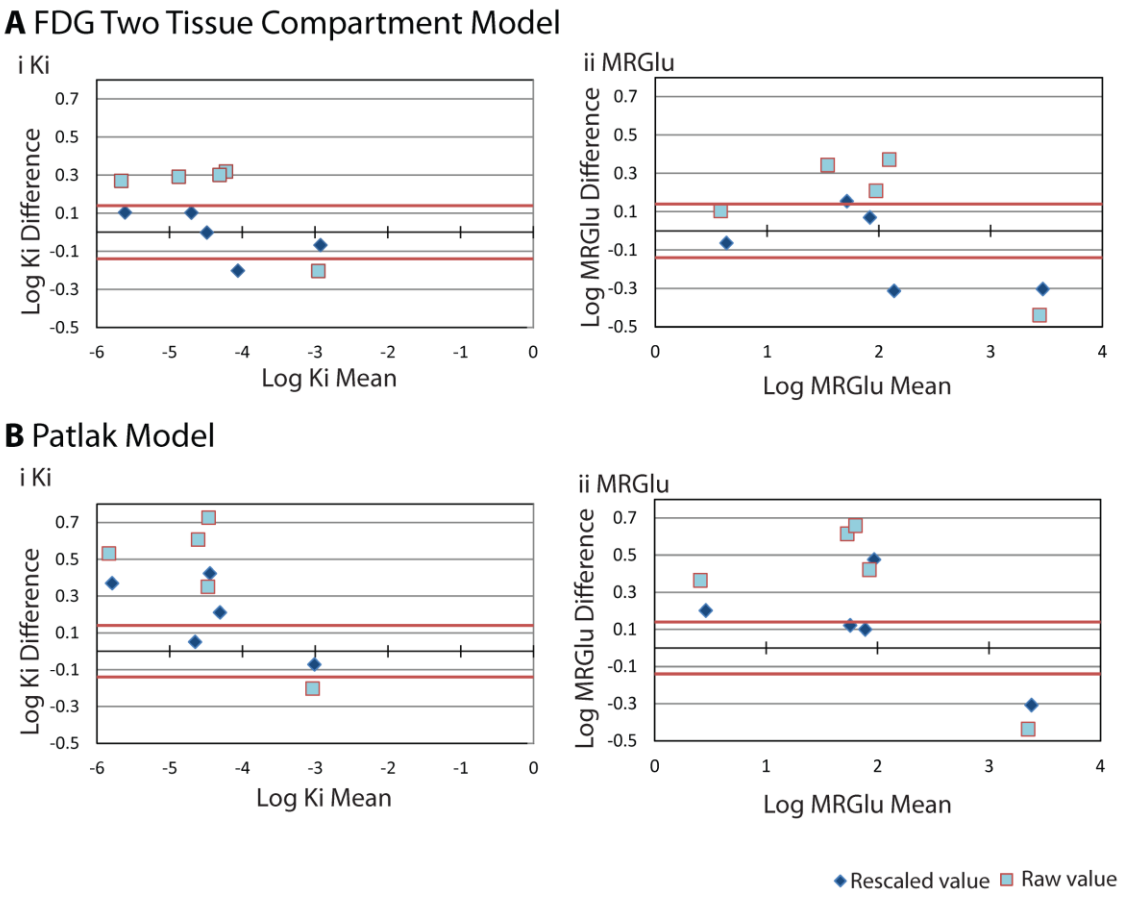


Figure 9-2 Bland Altman plots of log transformed tumour parameters

A. Bland Altman plots of log transformed measurements for raw and rescaled Ki (i) and MRGlu (ii) for breast lesions on the two baseline repeatability scans derived using the FDG two tissue compartment model. Lesions within -0.14 and 0.14 (dashed lines) indicate baseline scan measurements were within 15% of each other.

B. Bland Altman plots of log transformed measurements for raw and rescaled Ki (i) and MRGlu (ii) for breast lesions on the two baseline repeatability scans (Patlak model). Lesions within -0.14 and 0.14 (dashed lines) indicate baseline scan measurements were within 15% of each other.

9.3.2 PET response assessment using Ki and MRGlu

Based on the repeatability data, Ki and MRGlu values derived using rescaled 2TCM model only were considered for response evaluation. Table 9-2 reports the % change for each parameter at the post cycle 1 time point for the 7 of 8 evaluable breast lesions in relation to standard MRI RECIST and definitive RCB response.

Patient	Lesion	% change scaled Ki	% change scaled MRGlu	Hottest lesion SUV change (SUVmax, 90 minute scan)	Mid-MRI	EOT-MRI	RCB
1	**Breast**	-90.2	-90.6	-76.7	CR	CR	0
2	**Breast**	-50.2	-63.2	-39.9	SD	.	2
3	**Breast**	-69.0	-67.8	-52.5	PR	CR	0
4	Breast	-97.0	-97.7	-57.0	SD	PR	1
5	**Breast**	-66.6	-67.2	-46.9	PR	CR	0
6	**Breast**	-57.3	-62.0	-44.1	PR	PD	3
7	Breast	-	-	-44.5	SD	PD	3
8	Breast	-61.4	-66.2	-54.9	SD	CR	0
9	**Breast**	-89.4%	-88.4%	-72.4	PR	PR	0

Table 9-2 Dynamic response assessment for evaluable breast lesions

Per lesion % change in breast lesion Ki and MRGlu parameters. Where patients presented with more than one visible lesion, the SUV defined breast hottest lesion target is denoted in bold. Rescaled values only presented, thus participant 7 in whom no venous activity samples could obtained is considered non evaluable. Reference hottest lesion SUVmax change (90 minute scan acquisitions), mid and EOT MRI RECIST response and the definitive RCB response are shown for comparison. CR = complete response, SD = stable disease, PR = partial response, PD = progressive disease.

Mid- and EOT-MRI RECIST response was categorised as 'responder' (complete and partial) or 'non-responder' (stable or progression) for comparison with % change in Ki and MRGlu (Figure 9-3i and ii). No differences in distribution of % change in either parameter were present on visual inspection of box plots or statistically between categories of MRI response at either mid (p=1 for % change in both Ki and MRGlu) or EOT (p=0.13 for % change in Ki and MRGlu) time points. RCB score was categorised as 'responder' (RCB 0 or 1) or 'non responder' (RCB 2 or 3) for comparison with % Ki and MRGlu change. Visual inspection of the boxplots (Figure 9.3iii) suggests differences in the % change in Ki and MRGlu according to definitive RCB response category may be present (9-3 iii). This difference does not reach statistical significance (p=0.07)

for both dynamic parameters, Mann-Whitney U test), but is potentially an artefact of insufficient numbers for statistical validity within the current dataset. Full accrual will be required to definitively comment on possible predictive value of the dynamic PET parameters (Figure 9-3iii).

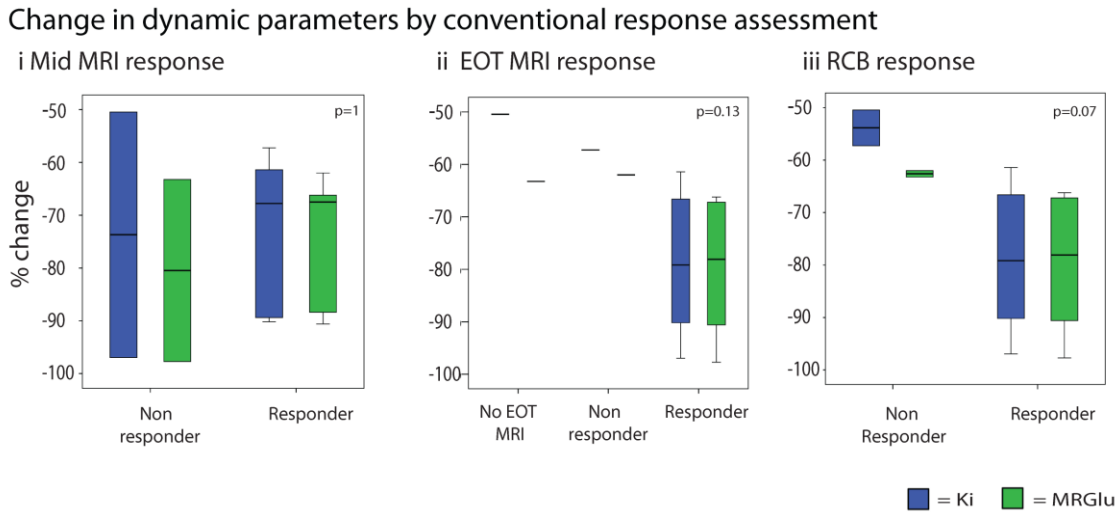


Figure 9-3 Distribution of dynamic parameter change according to conventional response assessment

Box plots showing comparison of distribution for Ki and MRGlu % change according to categories of (i) mid-MRI and (ii) EOT RECIST response and (iii) RCB response in n=8 breast lesions. No statistically significant difference in distribution of dynamic response parameter is present in the current dataset. Outlier patient 1 went on to achieve RCB0 pathological response, and outlier patient 6 RCB3.

The data presented in Chapter 8 suggest the predictive value of % change in SUV might be greatest when considering hottest lesion (breast or axilla) only as target for response assessment. The current dynamic breast dataset includes 6 patients where the evaluated breast lesion was also the SUV defined hottest lesion (Table 9.2). Considering distribution of % change in Ki and MRGlu between responding and non-responding categories within this subgroup only, no statistically significant differences in distribution of % change were present between categories of MRI response at mid (p=0.33 and 0.66 for % change in Ki and MRGlu respectively) or EOT (p= 0.17 for % change in both dynamic parameters) time points or between category of RCB response (p=0.13 for both dynamic parameters). For the EOT MRI and RCB comparisons this is potentially an artefact of insufficient numbers for statistical validity.

9.4 Segmented Dynamics

To facilitate assessment of SUV change over time the dynamic PET acquisitions obtained using the GE Discovery 710 64 slice PET-CT scanner (n=5) were reconstructed to provide four additional 10 minute 'static' views over the breast primary tumours at 40-50, 50-60, 60-70 and 70-80 minutes following tracer injection. SUV evaluation for each static reconstruction was performed using HERMES Hybrid Viewer Software version 1.4C Hermes as described previously (6.5.3) for comparison with the later 90, 120 and 180 minute static acquisitions.

Figure 9-4 presents the group mean SUVmax and SUVmean parameters for serial baseline and post cycle 1 acquisitions. Both parameters exhibit incremental increase with time following tracer injection that is steeper for the baseline acquisitions than following the first cycle of chemotherapy. Breast was hottest lesion target in only 2 of the 5 participants achieving RCB 0 (participant 5) and RCB3 (participant 6) respectively. It was not possible to derive static reconstructions from dynamic scan acquisitions performed using the GE Discovery VCT 64 slice PET-CT and consequently further datasets will be required to permit more meaningful exploration of the suitability of later image acquisition for response assessment or for consideration of hottest lesion targets on the relationship between SUV change and later pathological response. However given the increased baseline uptake on the later baseline acquisitions it is plausible that response assessment using the 120 or 180 minute acquisitions will be more clinically meaningful.

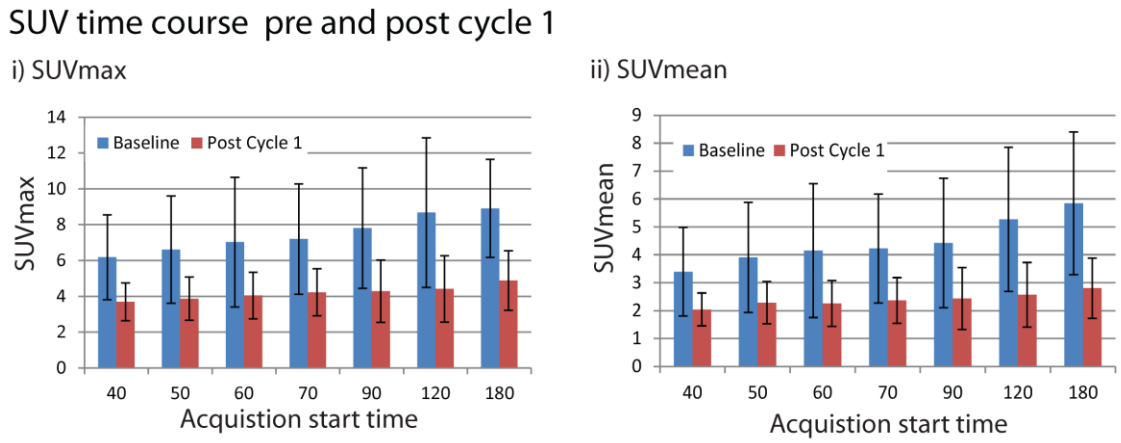


Figure 9-4 Change in SUV parameters with time

Mean baseline and post cycle 1 SUV parameters for the serial 10 minute static views acquired between 40 and 180 minutes following FDG tracer injection. The early static acquisitions were derived from summed dynamic data for the five participants imaged using the GE Discovery 710 64 slice PET-CT scanner. (i) SUVmax time course for FDG imaged breast tumours (ii)

SUVmean time course for FDG imaged breast tumours. Group mean and s.d. deviation shown for SUV parameters at each acquisition start time.

9.5 Conclusion

To our knowledge these are the first data reporting repeatability of dynamic parameters in FDG imaged breast tumours. Overall Ki and MRGlu values derived using the blood sample rescaled 2TCM model exhibited tighter repeatability than values derived from either model using only the image derived blood input; however both were outside +/-15% repeatability bounds. The FDG two tissue compartment model is the most commonly used PET metabolic model. This describes the biology of FDG exchange of the tracer between blood and tissue (tumour) compartments by fitting the tumour TAC to the standard two-tissue irreversible compartment model three rate constants, k1, k2 and k3 (FDG influx, efflux and phosphroylation respectively) plus blood volume fraction parameter (VB, vessel density). An alternative model, Patlak graphical analysis has been recommended by the EORTC (74) as the method of choice for a full quantitative analysis of dynamic FDG PET studies. This method is based on a linearization of the standard FDG model over a user defined interval, and assumes that the free concentration of FDG in tissue reaches a steady state and that binding/trapping is irreversible. With both methods the primary outcome parameters are K_i and MRGlu. However in the current dataset the Patlak model provided very poor reproducibility for breast lesions imaged using the FDG tracer even with rescaling. There are no published repeatability data in breast cancer for quantitative indices derived using FDG 2TCM or Patlak models, but mean percentage difference of 8%–10% has been reported in studies assessing reproducibility of Patlak Ki, derived from dynamic dataset comprising double imaged mixed primary and metastatic tumours (182). A trend for poorer reproducibility was observed with Ki but not SUV values in smaller tumours which was attributed to the impact of movement through the dynamic acquisition (182). Given the relatively small tumour volume in the current cohort (Table 8.1) which is representative of the neoadjuvant TNBC population this is likely to be a pertinent consideration in selecting the optimal reporting parameter for PET assessment neoadjuvant response.

To optimise patient acceptability and avoid the risks of invasive arterial blood sampling, an image-derived arterial input function was obtained using an ROI manually placed over the descending thoracic aorta to acquire the time course of tracer blood concentration. This method is widely accepted but risks contamination due to spill-over from other tissues, movement and partial volume effects compared to an arterial derived input function. To mitigate this, a venous activity sample at 60 minutes was secured where possible to permit

rescaling using the ratio of the manual blood sample to the time matched DA-ROI. The current data demonstrates the best repeatability for rescaled Ki evaluated using the 2TCM, not differing by more than 22% on the two scans (±0.2 on the logarithmic scale). However, although the shapes of the arterial blood clearance curves were similar between patients, the wide dispersal of the blood correction factor precludes generalisation where a venous sample could not be obtained.

Based on the repeatability data, Ki and MRGlu values derived using only rescaled 2TCM model were considered for response evaluation. In contrast to SUV change (Chapter 8), the current dataset provides no statistically significant difference in distribution of either dynamic response parameter between categories of MRI or RCB response. In addition to differential impact of patient movement on dynamic indices (182), discrepancies between SUV and full kinetic analysis results may be caused by changes in plasma glucose levels or differences in FDG plasma clearance among scans. No patient in the current cohort had serum glucose level greater than 6.2 mmol/l (normal range 4.4 to 6.1 mmol/L) and group mean serum glucose did not statistically differ between the scan visits (Chapter 8.4); therefore it seems unlikely that differences in blood glucose between scans explain the current data. Data in unselected breast cancer reports significant association between change in PET Ki and change in MRI tumour volume and later pathological response (76) and changes in K(1) and K(i) predicted both DFS and OS, whereas changes in SUV predicted OS only (77). In both studies PET response was assessed at mid or end of treatment but as SUV and Ki are not equivalent they may be differently affected by therapy induced metabolic changes with a different impact on the PET parameters at the post cycle 1 by treatment relative to later time points. The proportion of patients who had TNBC in these studies is unclear but is likely to be small, and differences in tumour biology of TNBC relative to other breast cancer phenotypes may therefore be relevant. There are few publications comparing SUV and 2TCM derived indices in cancer generally, however large differences between fractional change in SUV and Patlak derived MRGlu using an image derived input function have been reported for urological but not lung primary sites (181). Similarly in renal cancer comparison of on treatment change in tumor SUV with Patlak derived changes in Ki during serial PET imaging identified poor correlation between the two indices and the discrepancy was sufficient to result in conflicting conclusions regarding the progression of disease in some patients. Full accrual will be required to definitively comment on the predictive ability of Ki and MRGlu relative to SUV in TNBC however the emerging data suggest that these parameters may be less informative than hottest lesion SUV change.

It is disappointing that static reconstructions could not be derived from dynamic scan acquisitions performed using the GE Discovery VCT 64 slice PET-CT and further data will be

required for more robust assessment of the impact of acquisition interval on SUV in TNBC within the current study. However the current data clearly demonstrates steep incremental increase in SUVs with time following tracer injection that is most marked on the baseline compared to the cycle 1 scan. This is consistent with published data reporting diagnostic utility of FDG-PET unselected breast cancer (79, 80) and further supports requirement for consistent acquisition interval for comparison of pre and post treatment scans. SUVs should ideally be obtained after the tumor uptake reaches a plateau, and although earlier scan acquisitions facilitate patient throughput in a busy department, our observations suggest an uptake period longer than the 60-90 minute interval commonly applied for clinical imaging may be optimal. Using hottest lesion criteria we have demonstrated acceptable baseline repeatability to within 16%, corresponding to difference 0.15 on the log scale for SUVmax and SUVmean on the 120 minute acquisition and within 12% for SUVmax, 21% for SUVmean at 180 minutes (Chapter 7). Full accrual will be required to determine if delayed static scan acquisition beyond 90 minute confers a relatively greater magnitude of SUV change that is advantageous for differentiating categories of response in TNBC.

10 Conclusions

Given the expense and resource implication of imaging studies for researchers and participants it is imperative that early phase clinical trials select emerging tracers with the greatest likelihood of successfully addressing a clinical need. Despite prior promising data using a single (BT474) HER2+ breast cancer cell line (141) and anticipated rapid trajectory to first in man evaluation, the preclinical data presented in Chapter 5 demonstrates unexpected failure of a HER2 targeted DARPin radio-tracer designed to differentiate HER2 status of pre-clinical tumour xenografts models created using different breast cancer cell lines. This poor *in vivo* performance was disappointing but raises significant questions regarding suitability of a DARPin radiotracer for clinical evaluation as a HER2 diagnostic in its current form that may reduce the risk of proceeding to a futile and costly clinical study.

The tempo of accrual to the TNPET study was significantly slower than expected with many contributory factors including delays in the set up phase, issues with FLT production and the comparative rarity of TNBC despite the simple study design and use of established radiotracers. At the time of writing 12 of 18 recruited patients have competed all imaging research evaluations and recruitment to part B continues. Statistical validity of the analysed study data will be contingent on achieving full accrual and time only extension until Feb 2017 has been secured to allow adequate time for study completion. Strengths of the current PET data set include the clean TNBC phenotype and representation of the full spectrum of pathological response following neoadjuvant therapy. EANM control/quality assurance procedures concerning consistency in patient preparation scan procedure and image reconstruction for FDG image interpretation have been tightly adhered to (69) and consistency of scan scheduling in relation to therapy and acquisition time following tracer injection were all achieved.

Chapter 7 provides the first repeatability data for FDG-PET in breast cancer and for both tracers this is uniquely evaluated in a homogeneous population of patients with the TNBC biological subtype. Both tracers fulfilled per-patient repeatability criteria specified in the Part A analysis plan for at least one SUV measure. Considering all breast and nodal lesions the current data demonstrates that SUV intrinsic variability is 12-24% in both tracers, but is dependent on scan acquisition time and SUV parameter. PERCIST criteria (62) relating to group background activity at the 90 minute scan acquisition were met. Nevertheless a significant proportion of patient lesions were not evaluable using SUL and SUV peak measures, likely due to small

lesion size. As the recruited cohort are representative of the population with TNBC in whom neoadjuvant treatment would be recommended this strongly suggests that SUV peak parameters are impractical for response monitoring in this setting.

Following steering committee review of the repeatability data and recognising the impact of the unresolvable FLT production failures, the FDG tracer only was selected for progression to Part B. Subsequent to this decision two studies addressing questions concerning the predictive ability of FLT response assessment after 1 cycle of neoadjuvant chemotherapy have reported (94, 183). In both the recruited population comprised mixed breast cancer phenotypes and only limited numbers with TNBC. However neither baseline nor SUV change after 1 cycle predicted later pathological response irrespective of the SUV parameter selected. The larger data set noted a weak correlation between SUV change and pCR in the breast alone (183) but given the prognostic significance of residual nodal disease the clinical utility of this observation is unclear. It is disappointing that FLT could not continue to response assessment phase within the current study. Compared to FDG, FLT is more costly lack of commercial availability would hinder application for routine clinical use and with only four Part A patients imaged using this tracer further comment on the potential utility of this tracer in TNBC is not possible.

Defining pathological response following neoadjuvant chemotherapy using a minimum composite of residual breast and nodal disease to predict of long-term outcome is well established (25) and RCB continues to be the recommended international standard, delivering prognostic information for individual patients and also facilitating comparison of treatment outcomes within and across clinical trials (184, 185). In contrast the majority of publications reporting PET response have considered only the primary breast lesion, ignoring involved axillary disease (Appendix 11.1). PERCIST guidance recommends the % difference between the single most intense tumour on study 1 and study 2 is considered for response assessment (62) but no published data reporting neoadjuvant response in breast cancer has adopted this methodology. Two publications, subsequent to inception of the current study, pre-defined target lesion according to greatest baseline SUVmax, accepting that axilla rather than breast may be the index lesion for response evaluation (57, 75). Amongst 50 patients with TNBC, axillary nodal disease was target in 22% of patients (57). Using hottest lesion criteria to define the index tumour lesion in FDG imaged patients, the data presented in Chapter 7 demonstrates tighter repeatability of SUVmax and SUVmean parameters to within 12% and 16% at the 90 minute and 120 minute scan acquisitions and to within 12% for SUVmax, 21% for SUVmean at 180 minutes. Thus the 42% threshold reduction in SUVmax following 2 cycles

of neoadjuvant chemotherapy in TNBC suggested by Groheux *et al*. (57) comfortably exceeds these repeatability bounds. Using hottest lesion methodology to define target lesion for response assessment, the data presented in Chapter 8 suggests FDG PET will likely be informative of later pathological response at the even earlier post cycle 1 time point. It also seems likely discriminatory threshold for later RCB0/1 level of response will exceed the 15-30% level indicating unsuitability of the current PERCIST and EORTC guidance in this setting and the need for TNBC specific response criteria. However full accrual will be required for more definitive comment on this or the optimal scan acquisition time following tracer injection.

With the aspiration of future response adaptive therapy there remains considerable interest in identifying the optimal biomarker predictive of later pathological response and survival outcomes. Breast MRI provides better anatomical information than CT, and for prediction of RCB *volumetric* MRI imaging performs better than clinical (calliper) assessment of the breast primary particularly early in the course of neoadjuvant following 2 cycles of therapy (FTV2) (33). Nevertheless end of treatment imaging (FTV4) predicts RCB response with greater accuracy than either mid- or post cycle 2 time points (34) and multivariate analyses reported in the 2015 update identified equal contribution of MRI FTV2 and the histopathological variables (RCB class and tumor subtype defined by hormone and HER2 receptor status) for prediction of three year relapse free survival (33). MRI RECIST response at the mid-point of therapy was selected as the clean primary endpoint in the TNPET01 study but this poorly predicts later pathological response in the current dataset (Table 8.2). MRI measures including FTV do not consider nodal disease and, if the differential importance of nodal and breast disease between individuals with TNBC indicated by the current FDG PET data is confirmed, this may be pertinent to MRI methodology for neoadjuvant monitoring. Similarly in the interval from inception of the TNPET01 study, data demonstrating improvement in the predictive utility of RCB for survival outcomes by addition of parameters such post-treatment Ki-67 has been published (38, 186). However no tissue biomarker independently predictive of later pathological response or survival outcome has been identified and validated. If full study accrual confirms relationships between SUV change in hottest lesion target (breast or axilla site) and later RCB response it is plausible that some of the challenges associated with early tissue acquisition for response assessment may be avoided by ability to 'sample' the whole tumour burden using PET imaging alone or by using PET directed tissue acquisition to guide evaluation of early on-treatment biopsy derived changes.

10.1 Future Directions

Despite the slower than expected tempo of accrual the emerging data suggest the potential predictive utility of the FDG tracer at a more clinically desirable post cycle 1 stage justifying the completion of the current study. Specifically, full accrual will permit more definitive comment on the optimal scan methodology including acquisition interval, reporting parameter and target lesion definition using the widely available FDG tracer. Currently there are no TNBC specific tissue biomarkers that might be suitable for futureimaging application. However improved MRI visualisation of breast anatomy compared to CT might confer advantages for PET-MRI over PET-CT for neoadjuvant response evaluation in breast cancer. The limited data in breast cancer indicates equivalent performance in terms of qualitative lesion detection to PET-CT although significant differences in tracer uptake quantification was present, being most marked in normal lung, liver and muscle (187). In cancer imaging more generally PET-MRI performs as well as PET-CT for diagnostic purposes (188). However access to PET-MRI is limited and potential for integrating functional MRI parameters such as FTV with SUV for response evaluation remains unexplored.

In contrast to many other imaging study designs, the acquisition of research biopsy samples prior to and during the course of neoadjuvant and at definitive surgery provides an allied tissue resource permitting comparison with the imaging dataset. In addition, further work comparing tracer performance with biopsy derived indices of cell proliferation, apoptosis and glucose metabolism will be undertaken by my colleague Dr Sheeba Irshad (NIHR Clinical Lecturer, KCL) to better understand the molecular mechanisms driving the poor prognosis of TNBC patients using the tissue resource comprised of matched primary, on treatment and residual tissue following neoadjuvant chemotherapy that has been generated through this study.

11 Appendices

11.1 Overview of FDG response evaluation studies

Literature search updated March 2015

REF	n	Design	Eligibility Histology	Receptor data (if provided)	Therapy	PET			PET Parameter	PET response definition	Reference endpoint	Threshold change (%) for reference endpoint	Reference prediction
						response timing† (total)	Interval to static acquisition (min)	Normalisation					
(60)	59	P	HER2-	ER+: 61%	T	Week 2 (12 weeks)	60	LBM	SUL_{max}	52%	pCR	52%	Sn =75% Sp=74% NPV 89% PPV52%
(58)}	78	P	Unselected	Luminal:51% HER2+:15% TNEG:33% (IHC defined)	AT	1(3)	60	BW	SUV Ki-67	ROC	pCR (Honkoop A,B)	50% (SUV_{max})	Sn =85.7% Sp=60.9%
(59)	115	P	Unselected	Luminal:46% HER2+:32% TNEG:22% (IHC defined)	Mixed A/T/H and others	1(4-6)	80-90	BSA, glucose	SUV_{max}	ROC	pCR	75% (SUV_{max})	Sn =64% Sp=83% Accuracy 76%
(189)	54	P	HER2+	ER+ (74%)	T-H	1(	90	BSa, Glucose	SUV_{max}	ROC	pCR	60%	Sn =83% Sp=52%
(61)	61	P	HER- ER+		Mixed A ± T	1 (4-6)	80-90	BSA, glucose	SUV_{max}	REC	RFS	16%	Sn =92% Sp=79%
(83)	68	P	HER2+	ER+:56%	H +/- lapatinib	1	50		SUV_{max}	15% (EORTC)	pCR	15%	Sn =79% Sp=47%
(190)	60	P	Unselected	ER+:53% HER2+:20%	A-T sequential	1(6)	60	BW	SUV_{max}	15% (EORTC)	pCR (Sataloff)	15%	Sn =36% Sp=100%
(191)	32	P	Unselected	ER+: 78% HER2+: 22%	A-T sequential	**1**,4(8)	60	?	SUV_{max}	40%	pCR/near pCR	40%	Sn =63% Sp=92% NPV=91.6%
(192)	30	P	Unselected	-	A ±T	**1**, 5(8)	60	BSA	SUV_{max}	ROC	pCR	20%	Sn =90%

													Sp=74%
(193)	75	P	Unselected	-	A	**1**, 3 (6-8)	60	-	$SUV_{max,,}$ SUV_{mean} (42% isocontour)	ROC	Miller-Payne 4 & 5)	27% (SUV_{mean})	Sn =66% Sp=100%
(194)	47	P	Unselected	ER+: 60% HER2+: 34%	Mixed A, T & H regimens	1(4-6)	60	+/- BSA & glucose	$SUV_{max,,}$ $SUV_{max,-BSA-G}$	ROC	pCR (Sataloff, tumour & nodes)	60% $SUV_{max,,}$ $SUV_{max,-BSA-G}$	Sn =91% Sp=86%
									SUV_{av}			40%	Sn =82% Sp=89%
									$SUV_{av-BSA-G}$			50%	Sn =82% Sp=92%
(195)	107	P	Unselected	29% TNBC 24% HER2+ 47% ER+ HEr2-	Mixed, A and T	**1, 2** (8)	60		SUV_{max}				
(196)	37	R	Unselected	TNBC:27% ER+HER2-:46% ER+HER2+:16% ER-HER2+:11%	T-A sequential	**1** (4)	60	-	SUV_{max}	ROC	MRI response at end Taxane	>45% identified MRI responsers, <18% indicates non responders	
(71)	22	P	Unselected	-	A ±T	**1,** 2(3 or 4)	45	glucose	$SUV_{max,}$ SUV_{mean}	ROC	Honkoop A + B	55% (SUV_{max})	Cycle1 Sn =100% Sp=85%
													Cycle2 Sn =83% Sp=94%
(69)	64	P	Unselected	ER+: 56% HER2+: 23%	Mixed A-T or sequential	**1, 2**, 3,6(4-6)	60	BW, glucose	SUV_{max} USS, mammography	60,55,50 and thresholds	responders (Sataloff A+B)	60% (SUV_{max})	Cycle1 Sn =96% Sp=61%
													Cycle 2 Sn =95% Sp=89%
(70)	52	P	Nodal	ER+:63%	A or A-T	**1,2**,3,6(6)	60	Not stated -	SUV_{max}	ROC	pCR	Cycle 1	Sn =75%

			disease (unselected histology)	HER2+:19%	sequential						(Sataloff A &B in nodes)	50% (SUV_{max})	Sp=96% Accuracy=84%
												Cycle 2 50% (SUV_{max})	Sn =43% Sp=100% Accuracy=84%
(73)	69 & 81	P	Unselected	ER+: 63%	AT	**1,2**,(4-6)	45	Not stated	SUV_{max}	ROC	pCR/minimal residual disease	Cycle 1 45% (SUV_{max})	Sn =73% Sp=63% NPV 90%
												Cycle 2 55% (SUV_{max})	Sn =69% Sp=63% NPV 89%
(197)	40	P	Unselected	ER+: 48% HER2+: 30% TNBC: 23%	A-T sequential ±H	**2**	90	BW	SUVmax	ROC	RCB0-II vs. RCBIII	59.1%	Sn 68% Sp 75%
(198)	34	P	Unselected	ER+: 59% HER2+: 21% TNBC: 26%	A-T sequential	**2**, 4(6-8)	60	Not stated	SUV_{max}	45,50, 55%	(Miller-Payne 4 &5)	50% (SUV_{max})	Sn =100% Sp=30% NPV=100%
(199)	23	P	Unselected	Not Stated	A	2(6)	60	Not stated	- SUV_{max}	50%	Malignant cells ≤25% tumour area	50% (SUV_{max})	Sn =93% Sp=75%
(68)	64	P	ER+ HER2-	ER+ HER2-	A-T sequential	2 (8)	60	BW	TLG	ROC	Sartaloff response($T_{A\ or\ B}$; $N_{A,\ B\ \&\ C}$)	71%	Sn=74% Sp=89%
(57)	50	P	TNBC	TNBC	A-T sequential or A alone	2(8)	60	BW	$SUV_{max,}$	42%	pCR	42%	Sn=100% Sp=58%
(75)	30	P	HER2+	ER: 40%	A-T sequential	2(8)	60	BW	$SUV_{max,}$	SUV≤3	pCR	Uptake ≤3 on post cycle2 scan better than magnitude of change	Sn =86% Sp=94% Accuracy=90%
(32)	142	P	Unselected	ER+:70% HER2+:78%	A-T sequential ±H	2(8)	63	LBW	SUV_{max}	ROC	pCR	83.3% (SUV_{max})	Sn =67% Sp=96%
(200)	50		Unselected	Not stated	A ±	2(4-8)	60	Not stated	SUV_{max}	ROC	RECIST	40	Sn =77%

					Sequential T						(CR/PR vs PD/SD)	(SUV_{max})	Sp=80%
(72)	11	P	Unselected	-	A ±tamoxifen	1, 2, 3(3)	50	Not stated	SUV_{max}		Radiologyl or biopsy response		Significant early decrease in SUV in responders only
(77)	75	P (update, prior publications 08, 03)	Unselected	ER+: 55% HER2+:26%	A-T sequential	4(mean 6)	45 static dynamic	Not stated -	SUV_{peak} Kinetic-K1, Ki	ROC	pCR	(SUVpeak)	Kinetic parameter change more predictive
(85)	60	P	Unselected	ER+: 62% HER2+: 55%	A-T or T-A sequential	**4**, 8(8)	60 (NB 45-75minutes	BW	SUV_{max}	75%	pCR or >80% reduction RECIST	75%	Sn =78% Sp=60% ?sequence dependent
(201)	66	P	Unselected	ER+: 56% HER2+: 41%	T or A	End of tx (4)	60	Not stated	SUV_{peak}	ROC	pCR	84.8%	Sn =70% Sp=69.6%
(202)	41	P	Unselected	ER+: 46% HER2+: 24%	A or AT(H)	End of tx (mostly 3-4 cycle)	Not stated	Not stated	SUV_{peak}	50%	pCR	50%	Sn =86% Sp=38%
(159)	34	R	Unselected	Not stated	A-T sequential or TH	6(6)	Not stated	BW	SUV_{max}	ROC	pCR	63.9%	Sn =100% Sp=77.9%
(203)	50	P	Unselected	ER+: 52%	Mixed A T	End of tx (no of cycles not stated)	60	BW	SUV_{peak}	ROC	pCR	88%	Sn =100% Sp=56.5%
(204)	11	P	ER+		Endocrine (Leftrozole)	4(12) weeks	60	60	SUV_{max}	40%	pCR	40%	Metabolic response correlates with ki-67 change but not pCR

Key: Highlighted studies were those included in meta-analysis (55). Tumour receptor status classified according to oestrogen receptor positive or negative (ER+ or ER-); Human Epidermal Growth Factor Receptor 2 positive or negative (HER2 + or -) or TNBC if both ER and HER2 negative. For these studies ER- criteria were either undefined or <10% threshold.

Pathological response scoring systems: Miller Payne classifies pathological response according to a 5-step scale based on tumor cellularity in the excision/mastectomy specimen as compared with the pretreatment core biopsy; a score of 4 +5 defines those with >90% response or pCR(20). Sataloff separately

assess tumour response (A= total or near total therapeutic effect, B= >50% therapeutic effect) and nodal response (A +B= no nodal disease with or without evidence of therapy response) (22). Honkoop A = pCR, B =minimal residual disease

Chemotherapy regimen classified according to anthracycline (A), taxane (T) or Trastuzumab (H) based. †Where scans performed after multiple cycle sensitivity and specificity data refers to scan timepoint denoted in bold. Sn=sensitivity, Sp=specificity, NPV= negative predictive value, P=prospective, R= retrospective, BW=bodyweight, BSA=body surface area

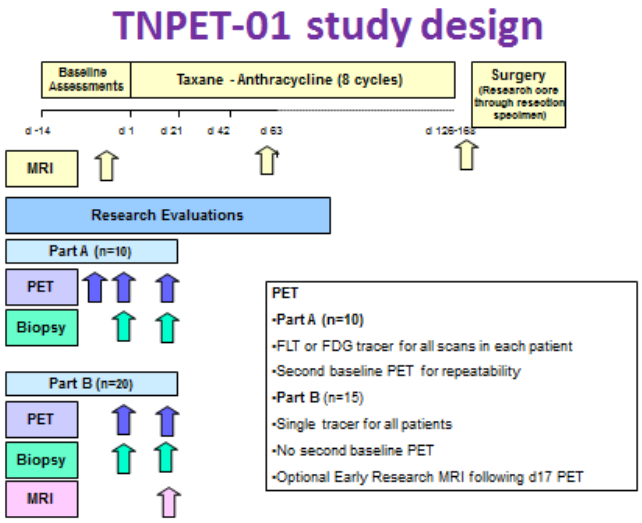
11.2 Overview of FLT breast cancer studies

Cohort	Receptor status	Therapy	FLT PET timing	Reference standard	PET parameter/endpoints	Conclusion	Ref
Response evaluation studies							
n=20 LABC and MBC	80% ER +	T (all pre-treated)	Post cycle 1 or 2	RECIST (at 3 or 4 cycles)	Predefined PET response 1. ($SUV_{60,av}$) of >20% 2. irreversible uptake (Ki) of >31%	1.SUV_{60},av: Sn=0.85, Sp= 0.80 2. median 40.2% decrease (responders) vs. 10.5% (non-responders). Ki reduction 51.1% (±28.4%, p<0.01)	(205)
n=13 LABC and MBC	Not stated	A	Post cycle 1	RECIST (at day 60)	1. FLT repeatability of SUV_{90} & Ki 2. SUV response	1. SD of mean % difference =10.5% (SUV90) and 15.1% (Ki); test-retest correlation coefficient ≥0.97 2. mean 41.3% (range 63.4 to 15.6%) & 52.9% (range 80.3 to 20.4%) decrease in SUV_{90} and ki respectively in RECIST responders. (No patients with progression)	(89)
n=14 MBC	Not stated	Ch or E	5/52	Tumour marker or imaging response (at mean 5.8 and 3.3 months respectively	SUVmean	1.mean FLT change correlates with later marker (r=0.79, p=0.001) and size (r=0.74, p=0.01) change	(90)
n=15 LABC	33% HER2+ 40% ER+ (60% TNBC)	Sequential A-T±H	Post cycle 1	Pathology (RCB)	SUVmax	Threshold drop ≥ 52.9% to define responders. Sn 83.3%, Sp 100%	(93)
n=17 LABC	30%TNBC 30% HER2+ 55% ER+	Mixed A +/- T	Post cycle 1	Pathology (pCR or not)	SUVmax, SUVmean	No predictive ability of absolute or SUV change	(94)
Mechanistic studies							
n=15 LABC	Not stated	A	Prior to cycle 2	No non PET reference	Comparison of simplified uptake measures (SUV) with full tracer kinetics (Non linear regression gold standard)	-ve bias for SUV change (no ki change corresponding to 11% decrease in SUV)	(146)
n=6 MBC	66% ER+ 33% HER2+	C	Post cycle 1	No non PET reference	Establish impact of TS inhibition on FLT tracer (SUV_{60} & ki)	Increased FLT retention in tumour but not normal tissue	(92)

Literature search (updated March 2015). E= endocrine therapy, Ch = mixed chemotherapy, unless classified according to anthracycline (A), taxane (T), Trastuzumab (H) or capecitibine (C) based regimen. MBC= metastatic breast cancer, LABC =locally advanced breast cancer. Sn=sensitivity, Sp=specificity, NPV= negative predictive value

11.3 Protocol Summary

Title of clinical trial	A randomised phase II trial of [18F]fluorothymidine and the standard tracer [18F]fluorodeoxyglucose in the assessment of systemic therapy response in triple negative breast cancer and their utility compared to conventional MRI imaging response, early ADC change and biopsy derived biomarkers
Protocol Short Title/Acronym	Triple negative breast cancer: novel functional imaging to determine early chemotherapy response/TNPET01
Trial Phase if not mentioned in title	-
Sponsor name	Kings College London/Guys and St Thomas' NHS Foundation Trust
Chief Investigator	Professor Andrew Tutt
Eudract number	2011-004220-34
REC number	11/L0/1492
Medical condition or disease under investigation	Imaging response biomarker study in triple negative breast cancer
Purpose of clinical trial	Phase II early imaging response biomarker study using PET-CT imaging for monitoring on treatment change in triple negative breast cancer
Primary objective	**Part A:** To confirm PET scan SUV measurement repeatability using [^{18}F]FDG and [^{18}F]FLT tracers **Part B:** To evaluate PET imaging using ^{18}F]FLT or [^{18}F]FDG) as methods for evaluating response to systemic therapy in primary triple negative breast cancer with respect to MRI response at 3 cycles
Secondary objective (s) **(Combined Part A and B data)**	Ascertain the optimal scan initiation time after [^{18}F]FLT and [^{18}F]FDG tracer administration in patients with triple negative breast cancer To correlate PET imaging response in breast and axillary lymph nodes with residual cancer burden (RCB) at definitive surgery To correlate PET imaging response using each tracer with blood and biopsy derived biomarkers. Non invasive assessment of Ki and k1 from this data set To obtain performance estimates for the ability of the Part B tracer (FDG or FLT) to report MRI response

	derived from integration of Apparent Diffusion Coefficient (ADC) and size change data at 3 cycles To obtain exploratory performance estimates for early MRI size and ADC evaluation on Diffusion Weighted MRI sequences after 1 cycle to report RECIST response at 3 cycles and RCB at definitive surgery To correlate MRI imaging ADC change with blood and biopsy derived proliferation biomarkers and apoptosis biomarkers To confirm the safety of [^{18}F]FLT in patients with breast cancer.
Trial Design	Single centre, non-therapeutic randomised open label Phase II trial with two parts
Trial Interventions	Part A 10 participants randomly allocated to FDGor FLTtracer will have two PET-CT scans separated by a minimum of 24 hours performed at baseline prior to chemotherapy. Participants will have a third scan at day 17±3 following the first cycle of chemotherapy to assess SUV response to treatment Part B 15 patients will be scanned once prior to commencing chemotherapy and again at day 14-21 post cycle 1 using the single tracer selected for progression to Part B according to end of Part A criteria . Optional study specific MRI scan performed at the end of cycle 1 (day 17±3) for early size change and apparent diffusion coefficient (ADC) evaluation. All particpants (A and B) will have a research core biopsy performed prior to chemotherapy, following their day 17±3 PET scan and through the definitive resection specimen at the time of surgery. All participants may have an optional research blood sample 
Endpoints	Part A (n=10) Final patient completing pre chemotherapy test-retest imaging in Part A.

	Tracers will be expected to achieve SUV repeatability of within ±15% and SUV reduction of 20-40% in at least 50% of MRI defined responders evaluable at the point the last patient is entered into part A. If these criteria are not met for a single tracer the alternative tracer will proceed to Part B of the study. If both tracers meet the criteria the tracer with the highest proportion of MRI defined responders with a drop in SUV of >20% will be selected to go through to part B. In the event of equal proportions the decision will be based on consensus between the team on which tracer performs the best overall. All consenting patients in part B will be followed using this single tracer. The study will terminate if neither tracer meets these criteria. Part B (n=15) On confirmation of tracer repeatability and after approval as a result of the Part A analysis the database will continue forward for a single tracer and Part A data contribute for Part B endpoint analysis. End of Study The end of the trial is the date of surgery of the last patient participating in the trial. This will be either completion of the last patients surgical visit if no IMP-related AE's have been seen of until any IMP-related AE monthly follow-up visits have been completed
Sample Size	25
Summary of eligibility criteria	Inclusion Female age 18 to 70 years Stage II-III biopsy proven early breast cancer for which primary chemotherapy is recommended. HER2 negative primary tumours (IHC 0 or 1+, or IHC 2+ and FISH non-amplified (ratio of *Her2* to chromosome 17 of more than 2.0) ER negative primary invasive breast cancer (Allred <3) ECOG PS of 0 or 1 Primary tumour size >2cm Eligible for neoadjuvant chemotherapy according to departmental protocols Able to comply with treatment plans, scheduled visits, all study PET imaging and biopsy procedures and follow-up Agree to use a medically acceptable birth control during the duration of their chemotherapy if of childbearing age. Exclusion Any prior treatment for the breast cancer Patients who are pregnant or breast feeding Evidence of metastatic disease at diagnosis precluding neoadjuvant chemotherapy.

	Requirement for concurrent radiotherapy treatment Serious medical condition or concurrent medical illness likely to compromise ability to complete chemotherapy course. Anticoagulation requirement which would preclude serial biopsy Diabetes Mellitus Any other problems that may make the patient unable to tolerate the PET scans or translational biopsies Investigational Medicinal Product in the previous 28 days
IMP, dosage and route of administration	Intravenous radiotracer administration at time 0 on day of PET imaging (3 sessions per patient Part A, 2 per patient Part B). Single IMP per patient [^{18}F]-fluorothymidine (FLT): maximum 200 MBq FLT: 6.5 mSv per administration The study IMP [^{18}F]-fluorothymidine is supplied by: The PET Imaging Centre St. Thomas' Hospital London SE1 7EH UK MIA(IMP) No. 11387 [^{18}F]-fluorodeoxyglucose (FDG): maximum 200 MBq of FDG: 4mSv per administration The study IMP [^{18}F]-FDG is MetaTrace FDG Solution for injection. Marketing authorisation number: PL 27150/001
Active comparator product(s)	Not applicable
Version and date of protocol amendments	Protocol version 7, 23rd July 2014

11.4 Summary of FLT productions issues (23/06/2014)

Problems underlying FLT production:

There has been considerable difficulty in producing research tracers consistently, which has increased over time due to an ageing cyclotron (22 years old; lifetime usually 10-15 years) and outdated radiochemistry infrastructure with limited hot cells available to enable multiple tracers to be produced. There was a flood in the cyclotron in 2013 and a subsequent prolonged period of downtime during May 2014. The cyclotron service was successfully recovered after both these events but further problems are likely and replacement parts for the cyclotron and industry support have not been commercially available for some time with all work done by engineers in house. This reliability/infrastructure issue will be addressed with the Phase II development of the PET Centre with new cyclotron and radiochemistry facilities anticipated for 2016/7.

There has also been underinvestment in staffing and a key member of staff responsible for FLT production and development left the centre in October 2013. A researcher from the Division of Imaging Sciences has been drafted in to support Radiochemistry development and troubleshooting of the ageing equipment but this single member of staff is the only individual in house with sufficient specialist skill to work in this area and is currently working on restarting FLT production. The issues with FLT production are not isolated – these have been in parallel with problems producing tracers for the clinical service in particular C11 methionine and C11 choline which over the past weeks has taken priority as well as other research tracers with projects and students dependent on them.

Attempts to address these:

To deal with the staffing issues, new posts were requested 3 months ago but only approved yesterday due to funding constraints. A contractor was also arranged as a short term position last week but that member of staff is unfortunately currently unable to work due to an acute illness. It is difficult to find people with appropriate expertise for this type of role.

Alternative sources for FLT have been looked into however there is only one commercial supplier of FLT with IMP standard but tracer is only produced once a week and therefore not able to be used for the final repeatability study.

A new more robust cassette based system for FLT production is planned but work to implement this to produce IMP quality FLT has not been possible due to lack of staff, who are currently fully stretched supporting the clinical service and research tracer production.

Ways forward:

In the short term work we will continue to address problems encountered with FLT production including the possibility of employing contract staff to assist and free up time for the existing member of the research team to restart FLT production.

In the medium term once new posts are in place the cassette based system should offer a more robust production.

In the longer term new facilities will address these problems.

Sally Barrington Tony Gee

23/06/2014

11.5 DARPin and Trastuzumab tracer preparation

11.5.1 ^{111}In DOTA conjugated $(HE)_3$-G3 DARPin (^{111}In DOTA-DARPin) tracer preparation

DOTA $(HE)_3$G3 DARPin was prepared and supplied for radiolabelling by Dr Robert Goldstein (UCL) (141). All radiolabelling was performed by Dr Margaret Cooper.

11.5.1.1 Radiolabelling of DARPin

To DOTA-G3 DARPin (50 µL, 15 µg) in 0.2M ammonium acetate, pH 6 (UCL, prepared by Rob Goldstein), was added 30 MBq (50 µL) ^{111}Indium chloride in 0.05M HCl (Covidien, Petten, Netherlands) and the radiolabelling reaction was heated at 37°C for 90 min. The radiolabelled DARPin was analysed by reverse phase HPLC using an Eclipse XDB-C8 5 µm, 4.6 x 150 mm column (Agilent) at 1 mL/min with a gradient of 0-60% B over 20 min, where solvent A is 0.1% TFA in water and solvent B 0.1% TFA in acetonitrile.

The radiolabelled antibody was diluted with 0.9% sodium chloride for injection to give 3 µg (6 MBq) ^{111}In-G3 DARPin per 80 µL injection volume. Standard solutions were prepared from the injection stock solution in order to calculate the % injected dose for biodistribution analysis.

11.5.1.2 Results

The radiolabelling efficiency was 92%.

In vivo experiment date	DARPin per mouse (µg)	Activity (MBq per injection)	Specific Activity (MBq/µg)
8.4.14	2.77	6.16	2.23
9.4.14	2.76	4.85	1.76
17.6.14	3	5.14	1.71

11.5.2 ^{111}In -CHX-A"-DTPA-trastuzumab tracer preparation

11.5.2.1 Conjugation

To 10 mg trastuzumab (Herceptin, Roche), which had been reconstituted to 21mg/mL according to the manufacturer's instructions, was added 50 mM EDTA (25 µL) to chelate any free metal in the antibody solution. After 30 min the solution was transferred to an ultracentrifugation tube (Vivaspin 15, 30,000 mwco, HY membrane) and the buffer was

exchanged for 0.1 M HEPES buffer, pH 8.9 by washing three times with HEPES buffer and concentrating the solution to approximately 1.5 mL during each centrifugation step. Trastuzumab in 0.1 M HEPES buffer was collected in a final volume of approximately 1.5 mL.

CHX-A”-DTPA bifunctional chelate (2 mg) in DMSO (40 µL) was added to the trastuzumab solution and the conjugation allowed to proceed at room temperature for 3 hr then overnight at 4°C.

Excess ligand was removed by ultracentrifugation as above and the buffer exchanged for 0.2 M ammonium acetate buffer, pH 6 by washing 8 times with ammonium acetate buffer and concentrating the solution to approximately 3 mL during each centrifugation step. The trastuzumab-conjugate was finally collected in approximately 0.7 mL 0.2 M ammonium acetate buffer, pH 6.

The concentration of the trastuzumab-conjugate was measured by Nanodrop 2000c UV-spectrometer (Thermo Scientific, USA) nm and calculated to be 13 mg/mL.

11.5.2.2 Radiolabelling

CHX-A”-DTPA-Herceptin (10µL, 130µg) was diluted to 1.86 mg/mL with 0.2 M ammonium acetate buffer, pH 6 (60µL). 111Indium chloride in 0.05M HCl (Covidien, Petten, Netherlands) was added (26 MBq, 33µL) and the radiolabelling reaction allowed to proceed at room temperature for 20 min. The radiolabelled antibody was analysed by size exclusion HPLC using a BioSep SEC-S-2000 column (Phenomenex, Macclesfield, UK) with an isocratic mobile phase of phosphate buffered saline, pH 7, containing 2 mM EDTA, and a flow rate of 1 mL/min. The retention time of the radioimmunoconjugate was 7 min and that of the unbound ^{111}In impurities was 10 min 30 sec and 12 min.

The radiolabelled antibody was diluted with 0.9% sodium chloride for injection to give 25µg (5 MBq) ^{111}In-CHX-A”-DTPA-trastuzumab per 80 µL injection volume. Standard solutions were prepared from the injection stock solution in order to calculate the %-injected dose for biodistribution analysis.

11.5.2.3 Results

The radiolabelling efficiency was 96.8% with less than 0.7% attributed to antibody aggregates.

12 References

1. CRUK. Cancer Statistics. 2012.
2. Chia SK, Speers CH, D'yachkova Y, Kang A, Malfair-Taylor S, Barnett J, et al. The imact of new chemotherapeutic and hormone agents on survival in a population-based cohort of women with metastatic breast cancer. Cancer. 2007;110(5):973-9.
3. Shigematsu H, Kawaguchi H, Nakamure Y, Tanaka K, Shiotani S, Koga C, et al. Significant surival improvment of patients with recurrrent breast cancer in the periods 2001-2008 and 1992-2000. BMC Cancer. 2011;31(11).
4. Dent RA, Trudeau M, Pritchard KI, Wedad MH, Kahn HK, Sawka CA, et al. Triple-Negative Breast Cancer: Clinical Features and Patterns of Recurrence. Clin Cancer Res. 2007;13:4429-34.
5. Liedtke C, Mazouni c, Hess KR, Andre F, Tordai A, Mejia JA, et al. Response to Neoadjuvant Therapy and Long-Term Survival in Patients with Triple-Negative Breast Cancer. J Clin Oncol. 2008;26(8):1275-81.
6. Kassam F, Enright K, Dent RA, Dranitsaris G, Myers J, Flynn C, et al. Survival outcomes for patients with metastatic triple-negative breast cancer: implications for clinical practice and trial design. Clin Breast Cancer. 2009;9(1):29-33.
7. Kallioniemi O-P, Holli K, Visakorpi T, Koivula T, Helin HH, Isola JJ. Association of C-erbB-2 protein over-expression with high rate of cell proliferation, increased risk of visceral metastasis and poor long-term survival in breast cancer. Int J Cancer. 1991;49(5):650-5.
8. Ross JS, Slodkowska EA, Symmans WF, Pusztai L, Ravdin PM, Hortobagyi GN. The HER-2 Receptor and Breast Cancer: Ten Years of Targeted Anti–HER-2 Therapy and Personalized Medicine. The Oncologist. 2009;14(4):320-68.
9. Singh JC, Jhaveri K, Esteva FJ. HER2-positive advanced breast cancer: optimizing patient outcomes and opportunities for drug development. Br J Cancer. 2014;111(10):1888-98.
10. Zidan J, Dashkovsky I, Stayerman C, Basher W, Cozacov C, Hadary A. Comparison of HER-2 overexpression in primary breast cancer and metastatic sites and its effect on biological targeting therapy of metastatic disease. Br J Cancer. 2005;93(5):552-6.
11. Botteri E, Disalvatore D, Curigliano G, Brollo J, Bagnardi V, Viale G, et al. Biopsy of liver metastasis for women with breast cancer: Impact on survival. The Breast. 2012;21(3):284-8.

12. Wolmark N, Wang J, Mamounas E, Bryant J, Fisher B. Preoperative Chemotherapy in Patients With Operable Breast Cancer: Nine-Year Results From National Surgical Adjuvant Breast and Bowel Project B-18. JNCI Monographs. 2001;2001(30):96-102.
13. Rastogi P, Anderson SJ, Bear HD, Geyer CE, Kahlenberg MS, Robidoux A, et al. Preoperative Chemotherapy: Updates of National Surgical Adjuvant Breast and Bowel Project Protocols B-18 and B-27. J Clin Oncol. 2008;26(5):778-85.
14. Kaufmann M, Von Minckwitz G, Mamounas E, Cameron D, Carey LA, Cristofanilli M, et al. Recommendations from an Internataional Consensus Conference on the Current Status and Future of Neoadjuvant Systemic Therapy in Primary Breast Cancer. Ann Surg Oncol. 2012;19(5):1508-16.
15. Cortazar P, Zhang L, Untch M, Mehta K, Costantino JP, Wolmark N, et al. Pathological complete response and long-term clinical benefit in breast cancer: the CTNeoBC pooled analysis. The Lancet. 2014;384(9938):164-72.
16. Symmans W, Wei C, Gould R, Zhang Y, Hunt K, Buchholz T, et al. Abstract S6-02: Long-term prognostic value of residual cancer burden (RCB) classification following neoadjuvant chemotherapy. Cancer Res. 2013;73(24 Supplement):S6-02.
17. von Minckwitz G, Untch M, Blohmer J-U, Costa SD, Eidtmann H, Fasching PA, et al. Definition and Impact of Pathologic Complete Response on Prognosis After Neoadjuvant Chemotherapy in Various Intrinsic Breast Cancer Subtypes. J Clin Oncol. 2012;30(15):1796-804.
18. Gianni L, Pienkowski T, Im Y-H, Roman L, Tseng L-M, Liu M-C, et al. Efficacy and safety of neoadjuvant pertuzumab and trastuzumab in women with locally advanced, inflammatory, or early HER2-positive breast cancer (NeoSphere): a randomised multicentre, open-label, phase 2 trial. The Lancet Oncology. 2012;13(1):25-32.
19. Earl HM, Vallier A-L, Hiller L, Fenwick N, Young J, Iddawela M, et al. Effects of the addition of gemcitabine, and paclitaxel-first sequencing, in neoadjuvant sequential epirubicin, cyclophosphamide, and paclitaxel for women with high-risk early breast cancer (Neo-tAnGo): an open-label, 2×2 factorial randomised phase 3 trial. The Lancet Oncology. 2014;15(2):201-12.
20. Ogston KN, Miller ID, Payne S, Hutcheon AW, Sarkar TK, Smith I, et al. A new histological grading system to assess response of breast cancers to primary chemotherapy: prognostic significance and survival. The Breast. 2003;12(5):320-7.
21. Symmans WF, Peintinger F, Hatzis C, Rajan R, Kuerer H, Valero V, et al. Measurement of Residual Breast Cancer Burden to Predict Survival After Neoadjuvant Chemotherapy. J Clin Oncol. 2007;25(28):4414-22.

22. Sataloff DM, Mason BA, Prestipino AJ, Seinige UL, Lieber CP, Baloch Z. Pathologic response to induction chemotherapy in locally advanced carcinoma of the breast: a determinant of outcome. J Am Coll Surg. 1995;180(3):297-306.
23. Chollet P, Amat S, Belembaogo E, Cure H, de Latour M, Dauplat J, et al. Is Nottingham prognostic index useful after induction chemotherapy in operable breast cancer? Br J Cancer. 2003;89(7):1185-91.
24. Mittendorf EA, Jeruss JS, Tucker SL, Kolli A, Newman LA, Gonzalez-Angulo AM, et al. Validation of a Novel Staging System for Disease-Specific Survival in Patients With Breast Cancer Treated With Neoadjuvant Chemotherapy. J Clin Oncol. 2011;29(15):1956-62.
25. Corben AD, Abi-Raad R, Popa I, Teo CHY, Macklin EA, Koerner FC, et al. Pathologic Response and Long-Term Follow-up in Breast Cancer Patients Treated With Neoadjuvant Chemotherapy: A Comparison Between Classifications and Their Practical Application. Arch Pathol Lab Med. 2013;137(8):1074-82.
26. Fumagalli D, Bedard PL, Nahleh Z, Michiels S, Sotiriou C, Loi S, et al. A common language in neoadjuvant breast cancer clinical trials: proposals for standard definitions and endpoints. The Lancet Oncology. 2012;13(6):e240-e8.
27. Hubbard R, Peintinger F, Liedtke C, Hatzis C, Kuerer HM, Valero V, et al. Prognostic value of residual cancer burden after neoadjuvant taxane-anthracycline chemotherapy in phenotypic subsets of breast cancer. J Clin Oncol. 2008;26(May 20 Suppl):abstract 544.
28. Carey LA, Dees EC, Sawyer L, Gatti L, Moore DT, Collichio F, et al. The Triple Negative Paradox: Primary Tumor Chemosensitivity of Breast Cancer Subtypes. Clin Cancer Res. 2007;13(8):2239-4.
29. Rouzier R, Perou CM, Symmans WF, Ibrahim N, Cristofanilli M, Andersone K, et al. Breast Cancer Molecular Subtypes Respond Differently to Preoperative Chemotherapy. Clin Cancer Res. 2005;11(16):5678-85.
30. Caudle AS, Gonzalez-Angulo AM, Hunt KK, Liu P, Pusztai L, Symmans WF, et al. Predictors of Tumor Progression During Neoadjuvant Chemotherapy in Breast Cancer. J Clin Oncol. 2010;28(11):1821-8.
31. Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, et al. New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228-47.

32. Tateishi U, Miyake M, Nagaoka T, Terauchi T, Kubota K, Kinoshita T, et al. Neoadjuvant Chemotherapy in Breast Cancer: Prediction of Pathologic Response with PET/CT and Dynamic Contrast-enhanced MR Imaging- Prospective Assessment 10.1148/radiol.12111177. Radiology. 2012;263(1):53-63.
33. Hylton NM, Gatsonis CA, Rosen MA, Lehman CD, Newitt DC, Partridge SC, et al. Neoadjuvant Chemotherapy for Breast Cancer: Functional Tumor Volume by MR Imaging Predicts Recurrence-free Survival—Results from the ACRIN 6657/CALGB 150007 I-SPY 1 TRIAL. Radiology. 2015:150013.
34. Hylton NM, Blume JD, Bernreuter WK, Pisano ED, Rosen MA, Morris EA, et al. Locally Advanced Breast Cancer: MR Imaging for Prediction of Response to Neoadjuvant Chemotherapy—Results from ACRIN 6657/I-SPY TRIAL. Radiology. 2012;263(3):663-72.
35. von Minckwitz G, Blohmer JU, Costa SD, Denkert C, Eidtmann H, Eiermann W, et al. Response-Guided Neoadjuvant Chemotherapy for Breast Cancer. J Clin Oncol. 2013;31(29):3623-30.
36. Yerushalmi R, Woods R, Hayes MM, Gelman KA. Ki67 in breast cancer: prognostic and predicitive potential. Lancet Oncol. 2010;11:174-83.
37. Jones RL, Salter J, A'Hern R, Nerurkar A, Parton M, Reis-Filho JS, et al. The prognostic significance of Ki67 before and after neoadjuvant chemotherapy in breast cancer. Breast Cancer Res Treat. 2009;116(1):53-68.
38. Sheri A, Smith IE, Johnston SR, A'Hern R, Nerurkar A, Jones RL, et al. Residual proliferative cancer burden to predict long-term outcome following neoadjuvant chemotherapy. Ann Oncol. 2015;26(1):75-80.
39. Connor AJM, Pinder SE, Elston CW, Bell JA, Wencyk P, Robertson JFR, et al. Intratumoural heterogeneity of proliferation in invasive breast carcinoma evaluated with MIBI antibody. The Breast. 1997;6(4):171-6.
40. Van Dort ME, Rehemtulla A, Ross BD. PET and SPECT Imaging of Tumor Biology: New Approaches towards Oncology Drug Discovery and Development. Curr Comput Aided Drug Des. 2008;4(1):46-53.
41. Younes M, Lechago LV, Somoano JR, Mosharaf M, Lechago J. Wide Expression of the Human Erythrocyte Glucose Transporter Glut1 in Human Cancers. Cancer Res. 1996;56(5):1164-7.
42. Been LB, Suurmeijer AJH, Cobben DC, Jager PL, Hoekstra HJ, Elsinga PH. [^{18}F]FLT-PET in oncology: current status and opportunities. Eur J Nucl Med Mol Imaging. 2004;31(12):1659-72.

43. Li D, Yao Q, Li L, Wang L, Chen J. Correlation between hybrid 18F-FDG PET/CT and apoptosis induced by neoadjuvant chemotherapy in breast cancer. Cancer Biol Ther. 2007;6(9):1442-8.
44. Buck A, Schirrmeister H, Kuhn T, Shen C, Kalker T, Kotzerke J, et al. FDG uptake in breast cancer: correlation with biological and clinical prognostic parameters. Eur J Nucl Med Mol Imaging. 2002;29(10):1317-23.
45. Basu S, Chen W, Tchou J, Mavi A, Cermik T, Czerniecki B, et al. Comparison of triple-negative and estrogen receptor-positive/progesterone receptor-positive/HER2-negative breast carcinoma using quantitative fluorine-18 fluorodeoxyglucose/positron emission tomography imaging parameters. Cancer. 2008;112(5):995-1000.
46. Tchou J, Sonnad SS, Bergey MR, Basu S, Tomaszewski J, Alavi A, et al. Degree of Tumor FDG Uptake Correlates with Proliferation Index in Triple Negative Breast Cancer. Mol Imaging Biol. 2010;12(6):657-62.
47. Straver ME, Aukema TS, Olmos RAV, Rutgers EJT, Gilhuijs KGA, Schot ME, et al. Feasibility of FDG PET/CT to monitor the response of axillary lymph node metastases to neoadjuvant chemotherapy in breast cancer patients. Eur J Nucl Med Mol Imaging. 2010;37(6):1069-76.
48. Groheux D, Giacchetti S, Moretti J-L, Porcher R, EspiÃ© M, Lehmann-Che J, et al. Correlation of high 18F-FDG uptake to clinical, pathological and biological prognostic factors in breast cancer. Eur J Nucl Med Mol Imaging. 2011;38(3):426-35.
49. Osborne JR, Port E, Gonen M, Doane A, Yeung H, Gerald W, et al. 18F-FDG PET of locally invasive breast cancer and association of estrogen receptor status with standardized uptake value: microarray and immunohistochemical analysis. J Nucl Med. 2010;51(4):543-50.
50. Kim BS, Sung SH. Usefulness of F-18-FDG uptake with clinicopathologic and immunohistochemical prognostic factors in breast cancer. Ann Nucl Med. 2012;26(2):175-83.
51. Brown RS, Goodman TM, Zasadny KR, Greenson JK, Wahl RL. Expression of hexokinase II and Glut-1 in untreated human breast cancer. Nucl Med Biol. 2002;29(4):443-53.
52. Bos R, van der Hoeven JJM, van der Wall E, van der Groep P, van Diest PJ, Comans EFI, et al. Biologic Correlates of 18Fluorodeoxyglucose Uptake in Human Breast Cancer Measured by Positron Emission Tomography. J Clin Oncol. 2002;20(2):379-87.
53. Specht JM, Kurland BF, Montgomery SK, Dunnwald LK, Doot RK, Gralow JR, et al. Tumor Metabolism and Blood Flow as Assessed by Positron Emission Tomography

Varies by Tumor Subtype in Locally Advanced Breast Cancer. Clin Cancer Res. 2010;16(10):2803-10.

54. Hussein YR, Bandyopadhyay S, Semaan A, Ahmed Q, Albashiti B, Jazaerly T, et al. Glut-1 Expression Correlates with Basal-like Breast Cancer. Transl Oncol. 2011;4(6):321-7.

55. Wang W, Zhang C, Liu J, Huang G. Is 18F-FDG PET accurate to predict neoadjuvant therapy response in breast cancer? A meta-analysis. Breast Cancer Res Treat. 2012;131:357-69.

56. Groheux D, Hindie E, Giacchetti S, Delord M, Hamy AS, de Roquancourt A, et al. Triple-negative breast cancer: early assessment with 18F-FDG PET/CT during neoadjuvant chemotherapy identifies patients who are unlikely to achieve a pathologic complete response and are at a high risk of early relapse. J Nucl Med. 2012;53(2):249-54.

57. Groheux D, Hindie E, Giacchetti S, Hamy A-S, Berger F, Merlet P, et al. Early assessment with F-18-fluorodeoxyglucose positron emission tomography/computed tomography can help predict the outcome of neoadjuvant chemotherapy in triple negative breast cancer. Eur J Cancer. 2014;50(11):1864-71.

58. Keam B, Im S-A, Koh Y, Han S-W, Oh D-Y, Cho N, et al. Early metabolic response using FDG PET/CT and molecular phenotypes of breast cancer treated with neoadjuvant chemotherapy. BMC Cancer. 2011;11.

59. Humbert O, Berriolo-Riedinger A, Riedinger JM, Coudert B, Arnould L, Cochet A, et al. Changes in 18F-FDG tumor metabolism after a first course of neoadjuvant chemotherapy in breast cancer: influence of tumor subtypes. Ann Oncol. 2012.

60. Connolly RM, Leal JP, Goetz MP, Zhang Z, Zhou XC, Jacobs LK, et al. TBCRC 008: Early Change in 18F-FDG Uptake on PET Predicts Response to Preoperative Systemic Therapy in Human Epidermal Growth Factor Receptor 2–Negative Primary Operable Breast Cancer. J Nucl Med. 2015;56(1):31-7.

61. Humbert O, Berriolo-Riedinger A, Cochet A, Gauthier M, Charon-Barra C, Guiu S, et al. Prognostic relevance at 5 years of the early monitoring of neoadjuvant chemotherapy using 18F-FDG PET in luminal HER2-negative breast cancer. Eur J Nucl Med Mol Imaging. 2014;41(3):416-27.

62. Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: Evolving Considerations for PET Response Criteria in Solid Tumours. J Nucl Med. 2009;50(Suppl 1):122S-50S.

63. Vanderhoek M, Perlman SB, Jeraj R. Impact of the Definition of Peak Standardized Uptake Value on Quantification of Treatment Response. J Nucl Med. 2012;53(1):4-11.

64. Daisne J-F, Duprez T, Weynand B, Lonneux M, Hamoir M, Reychler H, et al. Tumor Volume in Pharyngolaryngeal Squamous Cell Carcinoma: Comparison at CT, MR Imaging, and FDG PET and Validation with Surgical Specimen. Radiology. 2004;233(1):93-100.
65. Burri RJ, Rangaswamy B, Kostakoglu L, Hoch B, Genden EM, Som PM, et al. Correlation of positron emission tomography standard uptake value and pathologic specimen size in cancer of the head and neck. Int J Radiat Oncol Biol Phys. 2008;71(3):682-8.
66. Zhang Y, Hu J, Li J, Wang N, Li W, Zhou Y, et al. Comparison of imaging-based gross tumor volume and pathological volume determined by whole-mount serial sections in primary cervical cancer. Onco Targets Ther. 2013;6:917-23.
67. van Loon J, Siedschlag C, Stroom J, Blauwgeers H, van Suylen R-J, Knegjens J, et al. Microscopic Disease Extension in Three Dimensions for Non–Small-Cell Lung Cancer: Development of a Prediction Model Using Pathology-Validated Positron Emission Tomography and Computed Tomography Features. International Journal of Radiation Oncology*Biology*Physics. 2012;82(1):448-56.
68. Groheux D, Hatt M, Hindie E, Giacchetti S, de Cremoux P, Lehmann-Che J, et al. Estrogen receptor-positive/human epidermal growth factor receptor 2-negative breast tumors: early prediction of chemosensitivity with (18)F-fluorodeoxyglucose positron emission tomography/computed tomography during neoadjuvant chemotherapy. Cancer. 2013;119(11):1960-8.
69. Rousseau C, Devillers A, Sagan C, Ferrer L, Bridji B, Campion L, et al. Monitoring of Early Response to Neoadjuvant Chemotherapy in Stage II and III Breast Cancer by [18F]Fluorodeoxyglucose Positron Emission Tomography. J Clin Oncol. 2006;24(34):5366-72.
70. Rousseau C, Devillers A, Campone M, Campion L, Ferrer L, Sagan C, et al. FDG PET evaluation of early axillary lymph node response to neoadjuvant chemotherapy in stage II and III breast cancer patients. Eur J Nucl Med Mol Imaging. 2011;38(6):1029-36.
71. Schelling M, Avril N, Nahrig J, Kuhn W, Romer W, Sattler D, et al. Positron emission tomography using [(18)F]Fluorodeoxyglucose for monitoring primary chemotherapy in breast cancer. J Clin Oncol. 2000;18(8):1689-95.
72. Wahl RL, Zasadny K, Helvie M, Hutchins GD, Weber B, Cody R. Metabolic monitoring of breast cancer chemohormonotherapy using positron emission tomography: initial evaluation. J Clin Oncol. 1993;11(11):2101-11.

73. Schwarz-Dose J, Untch M, Tiling R, Sassen S, Mahner S, Kahlert S, et al. Monitoring Primary Systemic Therapy of Large and Locally Advanced Breast Cancer by Using Sequential Positoron Emission Tomography Imaging with [18F]Fluorodeoxyglucose. J Clin Oncol. 2009;27(4):535-41.
74. Young H, Baum R, Cremerius U, Herholz K, Hoekstra O, Lammertsma AA, et al. Measurement of clinical and subclinical tumour response using [18F]-fluorodeoxyglucose and positron emission tomography: review and 1999 EORTC recommendations. Eur J Cancer. 1999;35(13):1773-82.
75. Groheux D, Giacchetti S, Hatt M, Marty M, Vercellino L, de Roquancourt A, et al. HER2-overexpressing breast cancer: FDG uptake after two cycles of chemotherapy predicts the outcome of neoadjuvant treatment. Br J Cancer. 2013;109(5):1157-64.
76. Partridge SC, Vanantwerp RK, Doot RK, Chai X, Kurland BF, Eby PR, et al. Association between serial dynamic contrast-enhanced MRI and dynamic 18F-FDG PET measures in patients undergoing neoadjuvant chemotherapy for locally advanced breast cancer. J Magn Reson Imaging. 2010;32(5):1124-31.
77. Dunnwald LK, Doot RK, Specht JM, Gralow JR, Ellis GK, Livingston RB, et al. PET tumor metabolism in locally advanced breast cancer patients undergoing neoadjuvant chemotherapy: value of static versus kinetic measures of fluorodeoxyglucose uptake. Clin Cancer Res. 2011;17(8):2400-9.
78. Boellaard R, O'Doherty M, Weber W, Mottaghy F, Lonsdale M, Stroobants S, et al. FDG PET and PET/CT: EANM procedure guidelines for tumour PET imaging: version 1.0. Eur J Nucl Med Mol Imaging. 2010;37(1):181-200.
79. Beaulieu S, Kinahan P, Tseng J, Dunnwald LK, Schubert EK, Pham P, et al. SUV varies with time after injection in (18)F-FDG PET of breast cancer: characterization and method to adjust for time differences. J Nucl Med. 2003;44(7):1044-50.
80. Zytoon AA, Murakami K, El-Kholy MR, El-Shorbagy E. Dual time point FDG-PET/CT imaging... Potential tool for diagnosis of breast cancer. Clin Radiol. 2008;63(11):1213-27.
81. Boerner AR, Weckesser M, Herzog H, Schmitz T, Audretsch W, Nitz U, et al. Optimal scan time for fluorine-18 fluorodeoxyglucose positron emission tomography in breast cancer. Eur J Nucl Med. 1999;26(3):226-30.
82. Kumar R, Loving VA, Chauhan A, Zhuang H, Mitchell S, Alavi A. Potential of dual-time-point imaging to improve breast cancer diagnosis with (18)F-FDG PET. J Nucl Med. 2005;46(11):1819-24.

83. Gebhart G, Gámez C, Holmes E, Robles J, Garcia C, Cortés M, et al. 18F-FDG PET/CT for Early Prediction of Response to Neoadjuvant Lapatinib, Trastuzumab, and Their Combination in HER2-Positive Breast Cancer: Results from Neo-ALTTO. J Nucl Med. 2013;54(11):1862-8.
84. de Langen AJ, Vincent A, Velasquez LM, van Tinteren H, Boellaard R, Shankar LK, et al. Repeatability of 18F-FDG Uptake Measurements in Tumors: A Metaanalysis. J Nucl Med. 2012;53(5):701-8.
85. Schneider-Kolsky ME, Hart S, Fox J, Midolo P, Stuckey J, Hofman M, et al. The role of chemotherapeutic drugs in the evaluation of breast tumour response to chemotherapy using serial FDG-PET. Breast Cancer Res. 2010;12(3):R37.
86. Rasey JS, Grierson JR, Wiens LW, Kolb PD, Schwartz JL. Validation of FLT Uptake as a Measure of Thymidine Kinase-1 Activity in A549 Carcinoma Cells. J Nucl Med. 2002;43(9):1210-7.
87. Chalkidou A, Landau DB, Odell EW, Cornelius VR, O'Doherty MJ, Marsden PK. Correlation between Ki-67 immunohistochemistry and 18F-Fluorothymidine uptake in patients with cancer: A systematic review and meta-analysis. Eur J Cancer. (0).
88. Dowsett M, Nielsen TO, A'Hern R, Bartlett J, Coombes RC, Cuzick J, et al. Assessment of Ki67 in Breast Cancer: Recommendations from the International Ki67 in Breast Cancer Working Group. J Natl Cancer Inst. 2011;103(22):1656-64.
89. Kenny L, Coombes RC, Vigushin DM, Al-Nahhas A, Shousha S, Aboagye EO. Imaging early changes in proliferation at 1 week post chemotherapy: a pilot study in breast cancer patients with 3'-deoxy-3'[18F]fluorothymidine positron emission tomography. Eur J Nucl Med Mol Imaging. 2007;34:1339-47.
90. Pio BS, Park CK, Pietras R, Hsueh W, Satyamurthy N, Pegram MD, et al. Usefulnes of 3'-[F-18]Fluoro-3'-deoxythymidine with Positron Emission Tomography in Predicting Breast Cancer Response to Therapy. Mol Imaging and Biol. 2006;8:36-42.
91. Contractor KB, Kenny L, Stebbing J, Rosso L, Ahmad R, Jacob J, et al. [18F]-3'deoxy-3'-Fluorothymidine Positron Emission Tomography and Breast Cancer Response to Docetaxel. Clin Cancer Res. 2011: published online doi:10.1158/078-0432.CCR-11-783.
92. Kenny LM, Contractor KB, Stebbing J, Al-Nahhas A, Palmieri C, Shousha S, et al. Altered Tissue 3'-Deoxy-3'-[18F]Fluorothymidine Pharmacokinetics in Human Breast Cancer following Capecitabine Treatment Detected by Positron Emission Tomography. Clin Cancer Res. 2009;15(21):6649-57.
93. Crippa F, Agresti R, Sandri M, Mariani G, Padovano B, Alessi A, et al. 18F-FLT PET/CT as an imaging tool for early prediction of pathological response in patients with locally

advanced breast cancer treated with neoadjuvant chemotherapy: a pilot study. Eur J Nucl Med Mol Imaging. 2015:1-13.

94. Woolf DK, Beresford M, Li SP, Dowsett M, Sanghera B, Wong WL, et al. Evaluation of FLT-PET-CT as an imaging biomarker of proliferation in primary breast cancer. Br J Cancer. 2014;110(12):2847-54.

95. Moasser MM. The oncogene HER2: its signaling and transforming functions and its role in human cancer pathogenesis. Oncogene. 2007;26(45):6469-87.

96. Slamon DJ, Leyland-Jones B, Shak S, Fuchs H, Paton V, Bajamonde A, et al. Use of Chemotherapy plus a Monoclonal Antibody against HER2 for Metastatic Breast Cancer That Overexpresses HER2. N Engl J Med. 2001;344(11):783-92.

97. Marty M, Cognetti F, Maraninchi D, Snyder R, Mauriac L, Tubiana-Hulin M, et al. Randomized Phase II Trial of the Efficacy and Safety of Trastuzumab Combined With Docetaxel in Patients With Human Epidermal Growth Factor Receptor 2–Positive Metastatic Breast Cancer Administered As First-Line Treatment: The M77001 Study Group. J Clin Oncol. 2005;23(19):4265-74.

98. Moja L, Tagliabue L, Balduzzi S, Parmelli E, Pistotti V, Guarneri V, et al. Trastuzumab containing regimens for early breast cancer. The Cochrane database of systematic reviews. 2012;4:Cd006243.

99. Wolff AC, Hammond MEH, Hicks DG, Dowsett M, McShane LM, Allison KH, et al. Recommendations for Human Epidermal Growth Factor Receptor 2 Testing in Breast Cancer: American Society of Clinical Oncology/College of American Pathologists Clinical Practice Guideline Update. J Clin Oncol. 2013;31(31):3997-4013.

100. Geyer CE, Forster J, Lindquist D, Chan S, Romieu CG, Pienkowski T, et al. Lapatinib plus Capecitabine for HER2-Positive Advanced Breast Cancer. N Engl J Med. 2006;355(26):2733-43.

101. Baselga J, Cortés J, Kim S-B, Im S-A, Hegg R, Im Y-H, et al. Pertuzumab plus Trastuzumab plus Docetaxel for Metastatic Breast Cancer. N Engl J Med. 2012;366(2):109-19.

102. Swain SM, Kim S-B, Cortés J, Ro J, Semiglazov V, Campone M, et al. Pertuzumab, trastuzumab, and docetaxel for HER2-positive metastatic breast cancer (CLEOPATRA study): overall survival results from a randomised, double-blind, placebo-controlled, phase 3 study. The Lancet Oncology.14(6):461-71.

103. Verma S, Miles D, Gianni L, Krop IE, Welslau M, Baselga J, et al. Trastuzumab Emtansine for HER2-Positive Advanced Breast Cancer. N Engl J Med. 2012;367(19):1783-91.

104. Aurilio G, Disalvatore D, Pruneri G, Bagnardi V, Viale G, Curigliano G, et al. A meta-analysis of oestrogen receptor, progesterone receptor and human epidermal growth factor receptor 2 discordance between primary breast cancer and metastases. Eur J Cancer.50(2):277-89.

105. Bartlett AI, Starcyznski J, Robson T, MacLellan A, Campbell FM, van de Velde CJH, et al. Heterogeneous HER2 Gene Amplification: Impact on Patient Outcome and a Clinically Relevant Definition. Am J Clin Pathol. 2011;136(2):266-74.

106. Hoefnagel LDC, van der Groep P, van de Vijver MJ, Boers JE, Wesseling P, Wesseling J, et al. Discordance in ERα, PR and HER2 receptor status across different distant breast cancer metastases within the same patient. Ann Oncol. 2013;24(12):3017-23.

107. Niehans GA, Singleton TP, Dykoski D, Kiang DT. Stability of HER-2/neu Expression Over Time and at Multiple Metastatic Sites. J Natl Cancer Inst. 1993;85(15):1230-5.

108. Wang RE, Zhang Y, Tian L, Cai W, Cai J. Antibody-Based Imaging of HER-2: Moving into the Clinic. Curr Mol Med. 2013;13(10):1523-37.

109. Goldstein R, Sosabowski J, Vigor K, Chester K, Meyer T. Developments in single photon emission computed tomography and PET-based HER2 molecular imaging for breast cancer. Expert Rev Anticancer Ther. 2013;13(3):359-73.

110. Lub-de Hooge MN, Kosterink JG, Perik PJ, Nijnuis H, Tran L, Bart J, et al. Preclinical characterisation of 111In-DTPA-trastuzumab. Br J Pharmacol. 2004;143(1):99-106.

111. McLarty K, Cornelissen B, Scollard D, Done S, Chun K, Reilly R. Associations between the uptake of 111In-DTPA-trastuzumab, HER2 density and response to trastuzumab (Herceptin) in athymic mice bearing subcutaneous human tumour xenografts. Eur J Nucl Med Mol Imaging. 2009;36(1):81-93.

112. Dijkers EC, Kosterink JG, Rademaker AP, Perk LR, van Dongen GA, Bart J, et al. Development and characterization of clinical-grade 89Zr-trastuzumab for HER2/neu immunoPET imaging. J Nucl Med. 2009;50(6):974-81.

113. Oude Munnink TH, Korte MAd, Nagengast WB, Timmer-Bosscha H, Schröder CP, Jong JRd, et al. 89Zr-trastuzumab PET visualises HER2 downregulation by the HSP90 inhibitor NVP-AUY922 in a human tumour xenograft. Eur J Cancer.46(3):678-84.

114. McLarty K, Cornelissen B, Cai Z, Scollard DA, Costantini DL, Done SJ, et al. Micro-SPECT/CT with 111In-DTPA-pertuzumab sensitively detects trastuzumab-mediated HER2 downregulation and tumor response in athymic mice bearing MDA-MB-361 human breast cancer xenografts. J Nucl Med. 2009;50(8):1340-8.

115. Perik PJ, Lub-De Hooge MN, Gietema JA, van der Graaf WTA, de Korte MA, Jonkman S, et al. Indium-111–Labeled Trastuzumab Scintigraphy in Patients With Human

Epidermal Growth Factor Receptor 2–Positive Metastatic Breast Cancer. J Clin Oncol. 2006;24(15):2276-82.

116. Dijkers EC, Oude Munnink TH, Kosterink JG, Brouwers AH, Jager PL, de Jong JR, et al. Biodistribution of 89Zr-trastuzumab and PET Imaging of HER2-Positive Lesions in Patients With Metastatic Breast Cancer. Clin Pharmacol Ther. 2010;87(5):586-92.

117. Mortimer JE, Bading JR, Colcher DM, Conti PS, Frankel PH, Carroll MI, et al. Functional imaging of human epidermal growth factor receptor 2-positive metastatic breast cancer using (64)Cu-DOTA-trastuzumab PET. J Nucl Med. 2014;55(1):23-9.

118. Tamura K, Kurihara H, Yonemori K, Tsuda H, Suzuki J, Kono Y, et al. 64Cu-DOTA-trastuzumab PET imaging in patients with HER2-positive breast cancer. J Nucl Med. 2013;54(11):1869-75.

119. Wu AM. Engineered antibodies for molecular imaging of cancer. Methods (San Diego, Calif). 2014;65(1):139-47.

120. Spiridon CI, Guinn S, Vitetta ES. A comparison of the in vitro and in vivo activities of IgG and F(ab')2 fragments of a mixture of three monoclonal anti-Her-2 antibodies. Clin Cancer Res. 2004;10(10):3542-51.

121. Smith-Jones PM, Solit D, Afroze F, Rosen N, Larson SM. Early tumor response to Hsp90 therapy using HER2 PET: comparison with 18F-FDG PET. J Nucl Med. 2006;47(5):793-6.

122. Beylergil V, Morris PG, Smith-Jones PM, Modi S, Solit D, Hudis CA, et al. Pilot study of 68Ga-DOTA-F(ab')2-trastuzumab in patients with breast cancer. Nucl Med Commun. 2013;34(12):1157-65.

123. Robinson MK, Doss M, Shaller C, Narayanan D, Marks JD, Adler LP, et al. Quantitative immuno-positron emission tomography imaging of HER2-positive tumor xenografts with an iodine-124 labeled anti-HER2 diabody. Cancer Res. 2005;65(4):1471-8.

124. Reddy S, Shaller CC, Doss M, Shchaveleva I, Marks JD, Yu JQ, et al. Evaluation of the anti-HER2 C6.5 diabody as a PET radiotracer to monitor HER2 status and predict response to trastuzumab treatment. Clinical cancer research : an official journal of the American Association for Cancer Research. 2011;17(6):1509-20.

125. Olafsen T, Sirk SJ, Olma S, Shen CK, Wu AM. ImmunoPET using engineered antibody fragments: fluorine-18 labeled diabodies for same-day imaging. Tumour Biol. 2012;33(3):669-77.

126. Xavier C, Vaneycken I, D'huyvetter M, Heemskerk J, Keyaerts M, Vincke C, et al. Synthesis, Preclinical Validation, Dosimetry, and Toxicity of 68Ga-NOTA-Anti-HER2 Nanobodies for iPET Imaging of HER2 Receptor Expression in Cancer. J Nucl Med. 2013;54(5):776-84.

127. Lofblom J, Feldwisch J, Tolmachev V, Carlsson J, Stahl S, Frejd FY. Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications. FEBS Lett. 2010;584(12):2670-80.

128. Hoppmann S, Danikas A, Durrant C, Iveson P, Glaser M, Indrevoll B, et al. Abstract 353: *In vivo* PET imaging of HER2 expression with GE226: An ^{18}F-labelled affibody molecule. Cancer Res. 2012;72(8):supplement 1.

129. Orlova A, Tolmachev V, Pehrson R, Lindborg M, Tran T, Sandström M, et al. Synthetic Affibody Molecules: A Novel Class of Affinity Ligands for Molecular Imaging of HER2-Expressing Malignant Tumors. Cancer Res. 2007;67(5):2178-86.

130. Tolmachev V, Velikyan I, Sandstrom M, Orlova A. A HER2-binding Affibody molecule labelled with 68Ga for PET imaging: direct in vivo comparison with the 111In-labelled analogue. Eur J Nucl Med Mol Imaging. 2010;37(7):1356-67.

131. Kramer-Marek G, Kiesewetter DO, Capala J. Changes in HER2 Expression in Breast Cancer Xenografts After Therapy Can Be Quantified Using PET and 18F-Labeled Affibody Molecules. J Nucl Med. 2009;50(7):1131-9.

132. Kramer-Marek G, Bernardo M, Kiesewetter DO, Bagci U, Kuban M, Aras O, et al. PET of HER2-positive pulmonary metastases with 18F-ZHER2:342 affibody in a murine model of breast cancer: comparison with 18F-FDG.[Erratum appears in J Nucl Med. 2012 Jul;53(7):1169 Note: Omer, Aras [corrected to Aras, Omer]]. J Nucl Med. 2012;53(6):939-46.

133. Heskamp S, Laverman P, Rosik D, Boschetti F, van der Graaf WT, Oyen WJ, et al. Imaging of human epidermal growth factor receptor type 2 expression with 18F-labeled affibody molecule ZHER2:2395 in a mouse model for ovarian cancer. J Nucl Med. 2012;53(1):146-53.

134. Orlova A, Wallberg H, Stone-Elander S, Tolmachev V. On the selection of a tracer for PET imaging of HER2-expressing tumors: direct comparison of a 124I-labeled affibody molecule and trastuzumab in a murine xenograft model. J Nucl Med. 2009;50(3):417-25.

135. Baum RP, Prasad V, Müller D, Schuchardt C, Orlova A, Wennborg A, et al. Molecular Imaging of HER2-Expressing Malignant Tumors in Breast Cancer Patients Using Synthetic 111In- or 68Ga-Labeled Affibody Molecules. J Nucl Med. 2010;51(6):892-7.

136. Sorensen J, Sandberg D, Sandstrom M, Wennborg A, Feldwisch J, Tolmachev V, et al. First-in-human molecular imaging of HER2 expression in breast cancer metastases using the 111In-ABY-025 affibody molecule. J Nucl Med. 2014;55(5):730-5.

137. Lindberg H, Hofström C, Altai M, Honorvar H, Wållberg H, Orlova A, et al. Evaluation of a HER2-targeting affibody molecule combining an N-terminal HEHEHE-tag with a GGGC chelator for 99mTc-labelling at the C terminus. Tumor Biol. 2012;33(3):641-51.
138. Zahnd C, Kawe M, Stumpp MT, de Pasquale C, Tamaskovic R, Nagy-Davidescu G, et al. Efficient Tumor Targeting with High-Affinity Designed Ankyrin Repeat Proteins: Effects of Affinity and Molecular Size. Cancer Res. 2010;70(4):1595-605.
139. Zahnd C, Wyler E, Schwenk JM, Steiner D, Lawrence MC, McKern NM, et al. A Designed Ankyrin Repeat Protein Evolved to Picomolar Affinity to Her2. J Mol Biol. 2007;369(4):1015-28.
140. Goldstein R, Tolner B, Leyton J, Livanos M, Bhavsar G, Vigor K, et al. Abstract 3912: Pre-clinical developments of the G3 Designed ankyrin repeat protein (DARPin) for invivo assessment of HER2 expression Cancer Res. 2013;73(8):Supplement 1.
141. Goldstein R, Sosabowski J, Livanos M, Leyton J, Vigor K, Bhavsar G, et al. Development of the designed ankyrin repeat protein (DARPin) G3 for HER2 molecular imaging. Eur J Nucl Med Mol Imaging. 2015;42:288-301.
142. Kao J, Salari K, Bocanegra M, Choi Y, Girard L, Gandi J. Molecular Profiling of Breast Cancer Cell Lines Defines Relevant Tumor Models and Provides a Resource for Cancer Gene Discovery. PLoS One. 2009;4(7):e6146.
143. Behr TM, Sharkey RM, Sgouros G, Blumenthal RD, Dunn RM, Kolbert K, et al. Overcoming the nephrotoxicity of radiometal-labeled immunoconjugates: improved cancer therapy administered to a nude mouse model in relation to the internal radiation dosimetry. Cancer. 1997;80(12 Suppl):2591-610.
144. Mavi A, Urhan M, Yu JQ, Zhuang H, Houseni M, Cermik TF, et al. Dual Time Point 18F-FDG PET Imaging Detects Breast Cancer with High Sensitivity and Correlates Well with Histologic Subtypes. J Nucl Med. 2006;47(9):1440-6.
145. Lodge AM, Lucas DJ, Marsden KP, Cronin FB, O'Doherty JM, Smith AM. A PET study of 18FDG uptake in soft tissue masses. Eur J Nucl Med.26(1):22-30.
146. Lubberink M, Direcks W, Emmering J, van Tinteren H, Hoekstra OS, van der Hoeven JJ, et al. Validity of simplified 3'-deoxy-3'-[18F] fluorothymidine uptake measures for monitoring response to chemotherapy in locally advanced breast cancer. Mol Imaging Biol. 2012;14(6):777-82.
147. Menda Y, Ponto LLB, Schultz MK, Zamba GKD, Watkins GL, Bushnell DL, et al. Repeatability of gallium-68 DOTATOC positron emission tomographic imaging in neuroendocrine tumors. Pancreas. 2013;42(6):937-43.

148. Thie JA, Hubner KF, Smith GT. The Diagnostic Utility of the Lognormal Behavior of PET Standardized Uptake Values in Tumors. J Nucl Med. 2000;41(10):1664-72.
149. Martin Bland J, Altman D. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. The Lancet.327(8476):307-10.
150. Puhalla S, Mrozek E, Young D, Ottman S, McVey A, Kendra K, et al. Randomized Phase II Adjuvant Trial of Dose-Dense Docetaxel Before or After Doxorubicin Plus Cyclophosphamide in Axillary Node-Positive Breast Cancer. J Clin Oncol. 2008;26(10):1691-7.
151. Piedbois P, Serin D, Priou F, Laplaige P, Greget S, Angellier E, et al. Dose-dense adjuvant chemotherapy in node-positive breast cancer: Docetaxel followed by epirubicin/cyclophosphamide (T/EC), or the reverse sequence (EC/T), every 2 weeks, versus docetaxel, epirubicin and cyclophosphamide (TEC) every 3 weeks. AERO B03 randomized phase II study. Ann Oncol. 2007;18(1):52-7.
152. Miller KD, Soule SE, Calley C, Emerson RE, Hutchins GD, Kopecky K, et al. Randomized phase II trial of the anti-angiogenic potential of doxorubicin and docetaxel; primary chemotherapy as Biomarker Discovery Laboratory. Breast Cancer Res Treat. 2005;89(2):187-97.
153. Alvarez RH, Bianchini G, Hsu L, Cristofanilli M, J. Esteva FJ, Pusztai L, et al. The effect of different sequencing regimens of taxanes and anthracyclines in the primary systemic treatment (PST) of breast cancer (BC) patients (pts): M. D. Anderson Cancer Center retrospective analysis. J Clin Oncol. 2010;28(15s):abstract 548.
154. Guo B, Villeneuve D, Hembruff S, Kirwan A, Blais D, Bonin M, et al. Cross-Resistance Studies of Isogenic Drug-Resistant Breast Tumor Cell Lines Support Recent Clinical Evidence Suggesting that Sensitivity to Paclitaxel may be Strongly Compromised by Prior Doxorubicin Exposure. Breast Cancer Res Treat. 2004;85(1):31-51.
155. Li SP, Padhani AR. Tumor response assessments with diffusion and perfusion MRI. J Magn Reson Imaging. 2012;35(4):745-63.
156. Padhani AR, Liu G, Mu-koh D, Chenevert TL, Thoeny HC, Takahara T, et al. Diffusion-Weighted Magnetic Resonance Imaging as a Cancer Biomarker: Consensus and Recommendations. Neoplasia. 2009;11(2):102-25.
157. Yankeelov TE, Lepage M, Chakravarthy A, Broome EE, Niermann KJ, Kelley MC, et al. Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results. Magn Reson Imaging. 2007;25(1):1-13.

158. Rubesova E, Grell A-S, De Maertelaer V, Metens T, Chao S-L, Lemort M. Quantitative diffusion imaging in breast cancer: A clinical prospective study. J Magn Reson Imaging. 2006;24(2):319-24.

159. Park SH, Moon WK, Cho N, Chang JM, Im S-A, Park IA, et al. Comparison of diffusion-weighted MR imaging and FDG PET/CT to predict pathological complete response to neoadjuvant chemotherapy in patients with breast cancer. Eur Radiol. 2012;22(1):18-25.

160. Fangberget A, Nilsen L, Hole K, Holmen M, Engebraaten O, Naume B, et al. Neoadjuvant chemotherapy in breast cancer-response evaluation and prediction of response to treatment using dynamic contrast-enhanced and diffusion-weighted MR imaging. Eur Radiol. 2011;21(6):1188-99.

161. Sharma U, Danishad KKA, Seenu V, Jagannathan NR. Longitudinal study of the assessment by MRI and diffusion-weighted imaging of tumor response in patients with locally advanced breast cancer undergoing neoadjuvant chemotherapy. NMR Biomed. 2009;22(1):104-13.

162. Parikh J, Charles-Edwards G. Diffusion Weighted MRI of the Breast: Is there a role for apparent diffusion coefficient values in the prediction of response and in the early assessment of response to neoadjuvant chemotherapy? International Society for Magnetic Resonance in Medicine, 2011 meeting. 2011;abstract 1026.

163. Pickles MD, Gibbs P, Lowry M, Turnbull LW. Diffusion changes precede size reduction in neoadjuvant treatment of breast cancer. Magn Reson Imaging. 2006;24(7):843-7.

164. Karp JS, Surti S, Daube-Witherspoon ME, Muehllehner G. The benefit of time-of-flight in PET imaging: Experimental and clinical results. Journal of nuclear medicine : official publication, Society of Nuclear Medicine. 2008;49(3):462-70.

165. Surti S. Update on Time-of-Flight PET Imaging. J Nucl Med. 2015;56(1):98-105.

166. Li C-Y, Klohr S, Sadick H, Weiss C, Hoermann K, Schoenberg SO, et al. Effect of Time-of-Flight Technique on the Diagnostic Performance of 18F-FDG PET/CT for Assessment of Lymph Node Metastases in Head and Neck Squamous Cell Carcinoma. J Nucl Med Technol. 2014;42(3):181-7.

167. Koolen BB, van der Leij F, Vogel WV, Rutgers EJT, Vrancken Peeters M-JTFD, Elkhuizen PHM, et al. Accuracy of 18F-FDG PET/CT for primary tumor visualization and staging in T1 breast cancer. Acta Oncol. 2014;53(1):50-7.

168. Sabine VS, Faratian D, Kirkegaard-Clausen T, Bartlett JM. Validation of activated caspase-3 antibody staining as a marker of apoptosis in breast cancer. Histopathology. 2012;60(2):369-71.

169. Gonzalez MA, Tachibana KE, Chin SF, Callagy G, Madine MA, Vowler SL, et al. Geminin predicts adverse clinical outcome in breast cancer by reflecting cell-cycle progression. J Pathol. 2004;204(2):121-30.

170. Joshi S, Watkins J, Gazinska P, Brown JP, Gillett CE, Grigoriadis A, et al. Digital imaging in the immunohistochemical evaluation of the proliferation markers Ki67, MCM2 and Geminin, in early breast cancer, and their putative prognostic value. BMC Cancer. 2015;15:546.

171. Amos KD, Adamo B, Anders CK. Triple-negative breast cancer: an update on neoadjuvant clinical trials. Int J Breast Cancer. 2012;2012:385978.

172. Koolen BB, Pengel KE, Wesseling J, Vogel WV, Vrancken Peeters M-JTFD, Vincent AD, et al. FDG PET/CT during neoadjuvant chemotherapy may predict response in ER-positive/HER2-negative and triple negative, but not in HER2-positive breast cancer. The Breast. 2013;22(5):691-7.

173. Chang CC, Cho SF, Chen YW, Tu HP, Lin CY, Chang CS. SUV on dual-phase FDG PET/CT correlates with the Ki-67 proliferation index in patients with newly diagnosed non-Hodgkin lymphoma. Clin Nucl Med. 2012;37(8):e189-95.

174. Han B, Lin S, Yu LJ, Wang RZ, Wang YY. Correlation of (1)(8)F-FDG PET activity with expressions of survivin, Ki67, and CD34 in non-small-cell lung cancer. Nucl Med Commun. 2009;30(11):831-7.

175. Balko JM, Cook RS, Vaught DB, Kuba MG, Miller TW, Bhola NE, et al. Profiling of residual breast cancers after neoadjuvant chemotherapy identifies DUSP4 deficiency as a mechanism of drug resistance. Nat Med. 2012;18(7):1052-9.

176. Schmidt G, Meyberg-Solomayer G, Gerlinger C, Juhasz-Boss I, Herr D, Rody A, et al. Identification of prognostic different subgroups in triple negative breast cancer by Her2-neu protein expression. Arch Gynecol Obstet. 2014;290(6):1221-9.

177. Choi J, Kim DH, Jung WH, Koo JS. Metabolic interaction between cancer cells and stromal cells according to breast cancer molecular subtype. Breast Cancer Research : BCR. 2013;15(5):R78-R.

178. Munzone E, Botteri E, Sciandivasci A, Curigliano G, Nole F, Mastropasqua M, et al. Prognostic value of Ki-67 labeling index in patients with node-negative, triple-negative breast cancer. Breast Cancer Res Treat. 2012;134(1):277-82.

179. Balko JM, Giltnane JM, Wang K, Schwarz LJ, Young CD, Cook RS, et al. Molecular profiling of the residual disease of triple-negative breast cancers after neoadjuvant chemotherapy identifies actionable therapeutic targets. Cancer Discov. 2014;4(2):232-45.

180. Di Bonito M, Cantile M, Collina F, Scognamiglio G, Cerrone M, La Mantia E, et al. Overexpression of Cell Cycle Progression Inhibitor Geminin is Associated with Tumor Stem-Like Phenotype of Triple-Negative Breast Cancer. J Breast Cancer. 2012;15(2):162-71.
181. Cheebsumon P, Velasquez LM, Hoekstra CJ, Hayes W, Kloet RW, Hoetjes NJ, et al. Measuring response to therapy using FDG PET: semi-quantitative and full kinetic analysis. Eur J Nucl Med Mol Imaging. 2011;38(5):832-42.
182. Weber WA, Ziegler SI, Thödtmann R, Hanauske A-R, Schwaiger M. Reproducibility of Metabolic Measurements in Malignant Tumors Using FDG PET. J Nucl Med. 1999;40(11):1771-7.
183. Kostakoglu L, Duan F, Idowu MO, Jolles PR, Bear HD, Muzi M, et al. A Phase II Study of 3'-Deoxy-3'-18F-Fluorothymidine PET in the Assessment of Early Response of Breast Cancer to Neoadjuvant Chemotherapy: Results from ACRIN 6688. J Nucl Med. 2015;56(11):1681-9.
184. Bossuyt V, Provenzano E, Symmans WF, Boughey JC, Coles C, Curigliano G, et al. Recommendations for standardized pathological characterization of residual disease for neoadjuvant clinical trials of breast cancer by the BIG-NABCG collaboration. Ann Oncol. 2015;26(7):1280-91.
185. Provenzano E, Bossuyt V, Viale G, Cameron D, Badve S, Denkert C, et al. Standardization of pathologic evaluation and reporting of postneoadjuvant specimens in clinical trials of breast cancer: recommendations from an international working group. Mod Pathol. 2015;28(9):1185-201.
186. Esserman LJ, Berry DA, DeMichele A, Carey L, Davis SE, Buxton M, et al. Pathologic Complete Response Predicts Recurrence-Free Survival More Effectively by Cancer Subset: Results From the I-SPY 1 TRIAL—CALGB 150007/150012, ACRIN 6657. J Clin Oncol. 2012.
187. Pace L, Nicolai E, Luongo A, Aiello M, Catalano OA, Soricelli A, et al. Comparison of whole-body PET/CT and PET/MRI in breast cancer patients: lesion detection and quantitation of 18F-deoxyglucose uptake in lesions and in normal organ tissues. Eur J Radiol. 2014;83(2):289-96.
188. Spick C, Herrmann K, Czernin J. 18F-FDG PET/CT and PET/MRI Perform Equally Well in Cancer: Evidence from Studies on More Than 2,300 Patients. J Nucl Med. 2016;57(3):420-30.
189. Humbert O, Cochet A, Riedinger JM, Berriolo-Riedinger A, Arnould L, Coudert B, et al. HER2-positive breast cancer: 8F-FDG PET for early prediction of response to

trastuzumab plus taxane-based neoadjuvant chemotherapy. Eur J Nucl Med Mol Imaging. 2014;41(8):1525-33.

190. Kolesnikov-Gauthier H, Vanlemmens L, Baranzelli M-C, Vennin P, Servent V, Fournier C, et al. Predictive value of neoadjuvant chemotherapy failure in breast cancer using FDG–PET after the first course. Breast Cancer Res Treat. 2012;131(2):517-25.

191. Ueda S, Tsuda H, Saeki T, Osaki A, Shigekawa T, Ishida J, et al. Early reduction in standardized uptake value after one cycle of neoadjuvant chemotherapy measured by sequential FDG PET/CT is an independent predictor of pathological response of primary breast cancer. Breast Journal. 2010;16(6):660-2.

192. Smith IC, Welch AE, Hutcheon AW, Miller ID, Payne S, Chilcott F, et al. Positron emission tomography using [(18)F]-fluorodeoxy-D-glucose to predict the pathologic response of breast cancer to primary chemotherapy. J Clin Oncol. 2000;18(8):1676-88.

193. McDermott GM, Welch A, Staff RT, Gilbert FJ, Schweiger L, Semple SIK, et al. Monitoring primary breast cancer throughout chemotherapy using FDG-PET. Breast Cancer Res Treat. 2007;102:75-84.

194. Berriolo-Riedinger A, Touzery C, Riedinger J-M, Toubeau M, Coudert B, Arnould L, et al. [18F]FDG-PET predicts complete pathological response of breast cancer to neoadjuvant chemotherapy. Eur J Nucl Med Mol Imaging. 2007;34(12):1915-24.

195. Koolen BB, Pengel KE, Wesseling J, Vogel WV, Vrancken Peeters MJ, Vincent AD, et al. Sequential (18)F-FDG PET/CT for early prediction of complete pathological response in breast and axilla during neoadjuvant chemotherapy. Eur J Nucl Med Mol Imaging. 2014;41(1):32-40.

196. Hirakata T, Yanagita Y, Fujisawa T, Fujii T, Kinoshita T, Horikoshi H, et al. Early predictive value of non-response to docetaxel in neoadjuvant chemotherapy in breast cancer using 18F-FDG-PET. Anticancer Res. 2014;34(1):221-6.

197. Andrade WP, Lima EN, Osorio CA, do Socorro Maciel M, Baiocchi G, Bitencourt AG, et al. Can FDG-PET/CT predict early response to neoadjuvant chemotherapy in breast cancer? Eur J Surg Oncol. 2013;39(12):1358-63.

198. Martoni AA, Zamagni C, Quercia S, Rosati M, Cacciari N, Bernardi A, et al. Early (18)F-2-fluoro-2-deoxy-d-glucose positron emission tomography may identify a subset of patients with estrogen receptor-positive breast cancer who will not respond optimally to preoperative chemotherapy. Cancer. 2010;116(4):805-13.

199. Kumar A, Kumar R, Seenu V, Gupta S, Chawla M, Malhotra A, et al. The role of 18F-FDG PET/CT in evaluation of early response to neoadjuvant chemotherapy in patients with locally advanced breast cancer. Eur Radiol. 2009;19:1347-57.

200. Duch J, Fuster D, Munoz M, Fernandez PL, Paredes P, Fontanillas M, et al. 18F-FDG PET/CT for early prediction of response to neoadjuvant chemotherapy in breast cancer. Eur J Nucl Med Mol Imaging. 2009;36(10):1551-7.

201. Jung S-Y, Kim S-K, Nam B-H, Min SY, Lee SJ, Park C, et al. Prognostic Impact of [18F] FDG-PET in operable breast cancer treated with neoadjuvant chemotherapy. Ann Surg Oncol. 2010;17(1):247-53.

202. Choi JH, Lim HI, Lee SK, Kim WW, Kim SM, Cho E, et al. The role of PET CT to evaluate the response to neoadjuvant chemotherapy in advanced breast cancer: comparison with ultrasonography and magnetic resonance imaging. J Surg Oncol. 2010;102(5):392-7.

203. Kim SJ, Kim Sk, Lee ES, Ro J, Kang Sh. Predictive value of [18F]FDG PET for pathological response of breast cancer to neo-adjuvant chemotherapy. Ann Oncol. 2004;15(9):1352-7.

204. Ueda S, Tsuda H, Saeki T, Omata J, Osaki A, Shigekawa T, et al. Early metabolic response to neoadjuvant letrozole, measured by FDG PET/CT, is correlated with a decrease in the Ki67 labeling index in patients with hormone receptor-positive primary breast cancer: a pilot study. Breast Cancer. 2011;18(4):299-308.

205. Contractor KB, Kenny L, Stebbing J, Challapalli A, Al-Nahhas A, Palmieri C, et al. Biological basis of [^{11}C]choline-positron emission tomography in patients with breast cancer: comparison with [^{1}8F]fluorothymidine positron emission tomography. Nuc Med Commun. 2011;32(11):997-1004.